Peer Review Results for the Estrogen Receptor (ER) Binding Assay

Prepared for:

U.S. Environmental Protection Agency
Exposure Assessment Coordination and Policy Division
Office of Science Coordination and Policy
1200 Pennsylvania Avenue, N.W.
Washington, DC 20460

Prepared by:

Eastern Research Group, Inc.
14555 Avion Parkway
Suite 200
Chantilly, VA 20151-1102

13 April 2009

EPA Contract No. EP-C-07-059
Work Assignment 1-13

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

1.0 INTRODUCTION

In 1996, Congress passed the Food Quality Protection Act (FQPA) and amendments to the Safe Drinking Water Act (SDWA), which requires EPA to:

"…develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by naturally occurring estrogen, or other such endocrine effect as the Administrator may designate."

To assist the Agency in developing a pragmatic, scientifically defensible endocrine disruptor screening and testing strategy, the Agency convened the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC). Using EDSTAC (1998) recommendations as a starting point, EPA proposed an Endocrine Disruptor Screening Program (EDSP) consisting of a two-tier screening/testing program with in vitro and in vivo assays. Tier 1 screening assays will identify substances that have the potential to interact with the estrogen, androgen, or thyroid hormone systems using a battery of relatively short-term screening assays. The purpose of Tier 2 tests is to identify and establish a dose-response relationship for any adverse effects that might result from the interactions identified through the Tier 1 assays. The Tier 2 tests are multi-generational assays that will provide the Agency with more definitive testing data.

One of the test systems recommended by the EDSTAC for the Tier 1 battery was the estrogen receptor (ER) binding assay. Its purpose in the Tier-1 battery is to detect chemicals that may affect the endocrine system by binding to the ER. EPA requested the National Institute of Environmental Health Sciences (NIEHS) to prepare a comprehensive historical review and critical evaluation of ER binding assays that had been reported in the scientific literature. According to the Expert Panel convened by NIEHS which examined the review, no existing ER binding method was adequately detailed and standardized to be considered "validated". The Expert Panel recommended that the EPA focus its attention on validating an ER binding assay that uses recombinant receptor rather than receptor obtained from whole animals (rat uterine cytosol, RUC). The EPA subsequently initiated such a validation effort for a human recombinant ER (hrER) binding assay, but felt that it had already invested so much time and effort in validating the ER-RUC assay that it would be appropriate to complete the validation of that

assay as well. The current peer review focuses on the ER-RUC assay; peer review of the hrER assay will be conducted separately since validation of that assay is running somewhat behind that of the ER-RUC assay

Although the ER binding assay, like all of the assays in the Endocrine Disruptor Screening Program, was be peer reviewed separately from the other assays in the program, it is expected to be used in conjunction with other assays to determine the potential of a chemical to interact with the endocrine system. This "battery of assays" approach was designed by EDSTAC to take advantage of the unmediated sensitivity of *in vitro* assays and other pathway-specific assays, while at the same time incorporating whole-animal assays that can detect effects that may be due to metabolites and some of which can detect effects from several different endocrine-related pathways. A weight-of-evidence approach using information from all of the assays, not just the ER binding assay, will be used to determine whether a chemical substance has the potential to interact with the endocrine system.

The purpose of this peer review was to review and comment on the estrogen receptor (ER) binding assay for use within the EDSP to detect chemicals that may affect the endocrine system by binding to the estrogen receptor. The primary product peer reviewed for this assay was an Integrated Summary Report (ISR) that summarized and synthesized the information compiled from the validation process (i.e., detailed review papers, pre-validation studies, and inter-lab validation studies, with a major focus on inter-laboratory validation results). The ISR was prepared by EPA to facilitate the review of the assay; however, the peer review was of the validity of the assay itself and not specifically the ISR.

The remainder of this report is comprised of the unedited written comments submitted to ERG by the peer reviewers in response to the peer review charge (see Appendix A). Section 2.0 presents peer review comments organized by charge question, and Section 3.0 presents peer review comments organized by peer review expert. The Integrated Summary Report is presented in Appendix B and additional supporting materials are included in Appendix C.

The final peer review record for the ER Binding assay will include this peer review report consisting of the peer review comments, as well as documentation indicating how peer review comments were addressed by EPA, and the final EPA work product.

1.1 Peer Review Logistics

ERG initiated the peer review for the ER Binding assay on March 13, 2009. ERG held a pre-briefing conference call on March 23, 2009 to provide the peer reviewers with an opportunity to ask questions or receive clarification on the review materials or charge and to review the deliverable deadlines. Peer review comments were due to ERG on or before April 6, 2009.

1.2 Peer Review Experts

ERG researched potential reviewers through its proprietary consultant database; via Internet searches as needed; and by reviewing past files for related peer reviews or other tasks to identify potential candidates. ERG also considered several experts suggested by EPA. ERG contacted candidates to ascertain their qualifications, availability and interest in performing the work, and their conflict-of-interest (COI) status. ERG reviewed selected resumes, conflict-of-interest forms, and availability information to select a panel of experts that were qualified to conduct the review. ERG submitted a list of candidate reviewers to EPA to either (1) confirm that the candidates identified met the selection criteria (i.e., specific expertise required to conduct the assay) and that there were no COI concerns, or (2) provide comments back to ERG on any concerns regarding COI or reviewer expertise. If the latter, ERG considered EPA's concerns and as appropriate proposed substitute candidate(s). ERG then selected the five individuals who ERG determined to be the most qualified and available reviewers to conduct the peer review.

A list of the peer reviewers and a brief description of their qualifications is provided below.

- ***Patrick Balaguer, Ph.D.,*** focuses his research on nuclear receptors, endocrine disruptors and reporter gene technology. Dr. Balaguer is "Chargé de recherche" at the INSERM. He is the scientific leader of a group (8 persons), which belongs to a team directed by

Vincent Cavaillès "Hormonal Signaling, Environment and Cancer". This team belongs to INSERM U896 / IRCM (Cancer Research Institute of Montpellier).

Dr. Balaguer has more than 80 peer-reviewed publications. He has participated in the supervision of nine PhD students to the presentation of their thesis and supervised six post-doctoral students. Presently, he supervises three technicians, three PhD students and one post-doctoral student. Dr. Balaguer's scientific projects are focused on some major public health problems, such as hormone-dependent cancers (breast, prostate, ovary and endometrium), environmental endocrine disruptors and their interferences in human reproduction and carcinogenesis and finally hormone substitutive treatments for menopause symptoms and their consequences on cancer risk increase. Dr. Balaguer has strong experience in reporter cell lines for high throughput screening and for the detection of endocrine disrupters. His laboratory is well equipped for performing biochemistry, molecular biology, cellular biology, molecular modeling and has the equipment for luminescence detection (microplates luminometers, cooled and intensified CDD).

- ***Ingemar Pongratz, Ph.D.,*** is a Group Leader of a research group of students and post-doctorate students in the field of molecular biology with special interest in studies on endocrine disruption in the Department for Biosciences and Nutrition at the Karolinska Institute in Stockholm, Sweden. Dr. Pongratz obtained a degree in Biochemistry at the University of Stockholm 1989, a Ph. D. in Molecular Biology at the Karolinska Institutet, Department for Medical Nutrition in 1996. He was post-doc at the Karolinska Institutet, Department for Cell and Molecular Biology until 1999 and received an Assistant Professorship from the Swedish Research Council 2000-2004 and subsequently established his own group at the Department for Biosciences at Novum. Dr. Pongratz is Vice-Coordinator with practical and strategic responsibility for the CASCADE Network of Excellence, which is a European Union-funded research, training and dissemination network composed of 25 different research groups from 10 Countries. Dr. Pongratz has authored or co-authored several papers on dioxin mediated disruption of estrogen receptor transcriptional activity, on the role of the hsp90 chaperone complex on the AhR

transactivation process and identified the AhR partner factor ARNT as a estrogen receptor co-activator.

- ***Shlomo Sasson, Ph.D.,*** received his B.Sc. in Biology and Ph.D. in Medical Sciences (Biochemistry) from the Hebrew University of Jerusalem where he focused on the effects of androgenic steroids on rat thymus. He completed post-doctoral training at the School of Medicine at the University of Rochester with Dr. Angelo C. Notides focusing on the kinetic analyses of the interaction of antiestrogens with the estrogen receptor.

  In recent years Dr. Sasson became interested in diabetes research. His laboratory is engaged in several research projects investigating the regulation of the glucose transport system in skeletal muscles and in vascular endothelial and smooth muscles cells, studying the effects of lipid peroxidation on the function of insulin-secreting β-cells, and developing novel antidiabetic drugs

- ***Marie-Louise Scippo, Ph.D.,*** is a professor at the University of Liège and serves as head of the Laboratory of Food Analysis, Department of Food Sciences, Faculty of Veterinary Medicine. Dr. Scippo's research focuses on the development of chemical and biochemical methods of detection of residues and contaminants in food and various matrices. One of the research programs currently active in her lab is the detection of endocrine disruptors in various matrices using assays involving steroid receptor assays (both binding assays and transcriptional activation assays). Dr. Scippo has published in peer reviewed journals such as Analytica Chemica Acta, Analytical and Bioanalytical Chemistry, Accreditation and Quality Assurance, Talanta, Chromatographia, and The Analyst.

- ***William J. Welsh, Ph.D.,*** holds the title of Norman H. Edelman Professor in Bioinformatics and Computer-Aided Molecular Design in the Department of Pharmacology at the Robert Wood Johnson Medical School (RWJMS) in Piscataway NJ, University of Medicine and Dentistry of New Jersey (UMDNJ). Concurrently, he serves as Director of the UMDNJ Informatics Institute (http://informatics.umdnj.edu) that coordinates University-wide initiatives in bioinformatics, chemical informatics, and

computer-aided molecular design. In this role, Dr. Welsh also directs the GSBS Graduate Program is Bioinformatics. He is also PI and Director of the EPA-supported Environmental Bioinformatics and Computational Toxicology Center (ebCTC: http://www.ebCTC.org). He is a member of various schools, centers and institutes of excellence at UMDNJ and Rutgers University, including the Cancer Institute of New Jersey, the New Jersey Center for Biomaterials, Rutgers University School of Pharmacy, and the Environmental & Occupational Health Sciences Institute (EOHSI).

Dr. Welsh's laboratory specializes in the development and application of computational tools for predictive toxicology and pharmaceutical drug discovery. Dr. Welsh's publication record currently includes over 350 articles in peer-reviewed books and journals, over 600 abstracts from presentations at professional scientific meetings, and 28 patents and patent applications. He is the recipient of multiple awards and honors, including the Teacher of the Year Award (1983 and 1985), the St. Louis Research Award (1998), the University of Missouri-St. Louis Chancellor's Research and Creativity Award (2001), the University of Missouri Entrepreneur of the Year Award (2001), the Norman H. Edelman Endowed Professorship in Bioinformatics at UMDNJ-RWJMS (2003), and most recently the John C. Krantz, Jr. Award (2004). He serves on the advisory boards of several scientific journals. Dr. Welsh earned a B.S. degree in Chemistry from St. Joseph's University (Philadelphia, PA) in 1969 and a Ph.D. degree in Theoretical Physical Chemistry in 1974 from the University of Pennsylvania (Philadelphia, PA).

2.0 PEER REVIEW COMMENTS ORGANIZED BY CHARGE QUESTION

Peer review comments received for the ER Binding assay are presented in the sub-sections below and are organized by charge question (see Appendix A). **Peer review comments are presented in full, unedited text as received from each reviewer.**

2.1 Comment on the Clarity of the Stated Purpose of the Assay

Patrick Balaguer: Yes, the stated purpose of the assay is very clear. The purpose is to provide a test that enable to identify the potential of chemicals to interact with rat uterine estrogen receptors.

Ingemar Pongratz: Section 2, A.
In my opinion the test regarding the purpose of the assay should be modfied. The Estrogen Receptor Binding Assay is in itself only able to to identify compounds that bind to the estrogen receptors ERα and/or ERβ. However, compounds can interact with with the estrogen receptor system through alternative mechanisms that may not involve direct binding to the estrogen receptors.

Shlomo Sasson: The rat uterine cytosol estrogen receptor (RUC-ER) competitive binding assay is one in a battery of assays aimed at providing validated strategy for the screening of endocrine disruptors (e.g., xenobiotics and environmental chemicals) that interact with the ER in target tissues and alter their normal functions. The purpose of the RUC-ER binding assay in this extensive project is clearly stated. The background information on the ER binding assay and the various options and choices of ER preparations from different animal species and recombinant human ER are well-presented. The strengths and shortcomings of the RUC-ER binding assay in view of numerous previous studies were taken into account in selecting and including this assay in the collection of other cell-based- and *in vivo* assays.

Marie-Louise Scippo: Yes. It is clearly stated that the estrogen receptor rat uterus cytosol (ER-RUC) binding assay is one of the tests of a battery of complementary screens, included in the endocrine disruptor screening program Tier-1 battery.

It is also clearly stated that the aim of the ER-RUC is to detect an interaction with the estrogen receptor, not to identify the mechanism of action (stated i.e. p.8, 65, 69-70 of the ISR, and page 5 of Appendix 1).

What is less clear is the weight to give to the result obtained for an unknown chemical using the ER-RUC assay (interactive or not with the ER) within the battery of the Tier-1 program.

It should be interesting to give, in the introduction of the integrated summary report (ISR) (page 2, under C. "*The Tier1 battery of assays*"), a description of the strategy that will be used to classify a chemical as negative or positive after the Tier 1 screening, that includes various *in vitro* and *in vivo* assays (ISR, page 3, table 1), and to give the weight of each assay in the final decision of the Tier 1 screening.

William J. Welsh: The assay is described in a succinct and clear manner. The word "specificity", which is used on page 8 with respect to the saturation binding assay, should be defined in this context.

2.2 Comment on the Biological and Toxicological Relevance of the Assay as Related to its Stated Purpose

Patrick Balaguer: Yes, the assay is biologically and toxicologically relevant to the stated purpose. It enables to identify binder, equivocal or non binders. However, it does not allow to quantitative structure-activity relationship development for the moment.

Ingemar Pongratz: Section 2, B, C, D
Only partially. In my opinion the assay does not take into full account some issues.

Recent experiments have shown that one of the estrogen receptor isoforms namely ERβ is under circadian control in the mouse. The circadian system is well conserved so it therefore likely that this is occurs also in the rat model.

Depending on the timepoint of cytosol preparation, the levels of ERβ expression may be very low.

These low ERβ levels may result in cytosol preparations that fail to detect compounds that preferentially interact with the ER isoform and thus may considered "safe".

This point should be taken into account to avoid missinterpretation of obtained results.

A second point regading the biolgical relevance of the assay is the role of the molecular chaperone hsp90 and its role in maintaining the estrogen receptors in a ligand binding state. Previous work with hsp90 associated receptors, in particular receptors like the GR, has demonstrated an important role for the hsp90 complex. In the case of the estrogen receptor ERα experiments have demonstrates that hsp90 is important but not crucial, for ligand binding. I am concerned that during cytosol preparation, the hsp90 complex may dissociate which would negatively impact on the receptors ligand binding activity. The presence of molybdate will stabilize the complex but it may not be sufficient.

In the case of the second estrogen receptor isoform ERβ, very little is know regarding the putative role of hsp90 and regarding the stability of the complex. Again this may cause problems and in particular may explain some of the interlaboratory variation.

Shlomo Sasson: The stated purpose of the RUC-ER binding assay is clear- it aims at identifying compounds that interact with the ER by testing their potential to compete with the natural ligand, 17β-estradiol, for binding with rat ER. It is clearly acknowledged that this competitive binding assay is not aimed at ascertaining the functional properties (weak/strong agonists, partial/mixed agonists or antagonists) of the test chemicals. However, the goal to categorize these molecules according to their binding affinities with the receptor is feasible and suits the formal objectives of the assay. It should be noted, however, that while there is a high degree of confidence that high affinity binders/competitors may activate or inhibit ER function in *in vitro* and *in vivo* models, it is doubtful whether weak or very weak binders/competitors would substantially interact with the ER and exert biological functions *in vivo*. Thus, this RUC-ER assay (and most likely the hER assay that is evaluated in parallel) allows the classification of potential endocrine disruptors by virtue of their intrinsic binding affinity for the ER. Further analysis of the

biological or toxicological effects of these compounds entails independent cell- and animal-based assays.

Marie-Louise Scippo: Yes, it is, even if this single assay gives no indication about the toxicity of a chemical.

The stated purpose of assays involving estrogen receptors is to evidence an estrogenic or an anti-estrogenic activity of the test chemicals.

For both estrogenic and anti-estrogenic compounds, in most cases, the first step of their biological activity is the binding to the estrogen receptor (ER), before target genes transcriptional activation (estrogenic compounds) or inhibition of it (anti-estrogenic compounds). There are two isoforms α and β of the estrogen receptor, and some compounds act more specifically on one or another form.

As the rat uterus tissue displays both α and β isoforms, the ER-RUC allows detecting both α and β ER ligands, which is an advantage over the binding assays using a recombinant receptor of a single isoform.

As the transcriptional activation assay of the Tier-1 battery is specific to the α isoform of the receptor, it is relevant to use the ER-RUC allowing the detection of the binding to both α and β isoforms, in a screening approach using a battery of complementary assays, in order to decrease the rate of false positive results of the Tier-1 testing.

William J. Welsh: The assay is relevant to the stated purpose, i.e., to determine the ability of a compound to interact with the ERs isolated from rat uteri". As discussed in the ISR on page 64, it is not clear whether testing compounds at 1 milliMolar (1 mM) is toxicologically relevant. Although the rationale for testing compounds at this high, physiologically non-relevant concentration is explained adequately (pp. 64-65), the results obtained may be difficult to explain and subject to mis-interpretation and even false conclusions.

2.3 Provide Comments on the Clarity and Conciseness of the Protocol in Describing the Methodology of the Assay such that the Laboratory can a) Comprehend the Objective, b) Conduct the Assay, c) Observe and Measure Prescribed Endpoints, d) Compile and Prepare Data for Statistical Analyses, and e) Report Results

Ingemar Pongratz: Section 2F

The protocol is technically sounds, and very detailed. In my opinion efforts should be done to make it easier to follow. It is stated that the groups performing the assays should have a suitable background in the field so they are well informed about the method. However I feel that the protocol is difficult to follow. In general I feel that an "Excecutive Summary" would be beneficial in particular to understand the purpose of the assay.

This comment is also relevant for point b; how to conduct the assay. Again a short annex would be improve clarity.

I have very limited experience regarding statistical analysis so I feel I am not competent to provide input regading this point.

Marie-Louise Scippo: The answer to questions a) to e) is globally YES, but some clarifications are needed (see here below).

William J. Welsh: The protocol adequately describes the assay in a clear and concise manner. In terms of advice, a *glossary of terms* would facilitate clarity between the EPA and the laboratory and to help avoid any misunderstanding or mis-interpretation in meaning of specific terms described in the protocol(s).

2.3.1 Comprehend the Objective

Patrick Balaguer: Yes, the description of the methodology is clear and concise and enable the comprehension of the objective.

However the finding that strong binders were considered « equivocal » rather than clearly interactive in one laboratory when tested blindy deserves and need to be noted in the protocol.

When binding at the concentrations tested showed clear interaction with the receptor, the laboratory did adjust the test concentration range to a more dilute range that would allow characterization of the full binding curve as required by the protocol. Thus, the protocol need to be adjusted to emphasize that a full curve must be obtained where there are clear indications of binding.

Shlomo Sasson: The objective of the RUC-ER binding assay is clear. The background information on the assay and its inclusion in the study are well-discussed. The assay was developed following an in-depth review of the literature, which provides the scientific basis for the assay. The standardization and attempts to optimize the assay following the First Interlaboratory Study resulted in detailed, comprehensive and clear protocols. Experienced laboratories, such as the 3 that participated in the Second Interlaboratory Study, are expected to adapt and conduct the assay with minimal deviations from the protocols. Nevertheless, the results obtained from these 3 laboratories and their comments on problems encountered in the course of the study do indicate the need for further examination and optimization of the protocols.

2.3.2 Conduct the Assay

Patrick Balaguer: Yes, the description of the methodology is clear and concise and enable the conduction of the objective;
Again, the protocol need to be adjusted to emphasize that a full curve must be obtained where there are clear indications of binding.

Shlomo Sasson: The conduct of the assay is expected to follow the detailed protocols. The two-test approach (the [^{3}H]-17β-estradiol saturation binding analysis, on one hand, and the competitive binding assay, on the other) is very important as it provides intra- and inter-laboratory quality controls and intrinsic assay controls. There are, however, certain problems that were encountered 'on the bench' or when data were collected, analyzed and interpreted. These issues are presented and discussed at the end of this section under **f**.

2.3.3 Observe and Measure Prescribed Endpoints

Patrick Balaguer: Yes, the description of the methodology is clear and concise and enable the observation and the measurement of prescribed endpoints;
Again, the protocol need to be adjusted to emphasize that if binding is clearly observed at 100 pM-1mM, the laboratory did adjust the test concentration range to a more dilute range.

Shlomo Sasson: The protocols are clear, detailed and meticulously describe all necessary technical aspects- from uteri excision to data acquisition and analyses. By conducting the assay according to protocols each laboratory is expected to observe and measure prescribed endpoints. In some cases along the Second Interlaboratory Validation Assay not all endpoints were achieved- mostly due to deviations from the prescribed protocols and/or due to technical problems.

2.3.4 Compile and Prepare Data for Statistical Analyses

Patrick Balaguer: Yes, the description of the methodology is clear and concise and enable the compilation and the preparation of the data or the statiscal analysis and the report of the results;

Shlomo Sasson: The platform provided to compile and prepare the data for kinetic analyses followed by statistical analyses has served well the participating laboratories and suits the requirements of the assay. Not being an expert in statistics, I cannot comment on the statistical methodology (Appendices 4 and 9). However, Appendix 13, which depicts graphs of acceptable saturation and competitive binding curves, raises some reservations on the acceptability of these graphs: the range of 17β-estradiol IC_{50} values (TRL's and Hammer's curves, Appendix 13/pp. 2-3) and of the weak positive runs (*ibid*, pp. 5-6) covers nearly an entire order of magnitude. It appears that narrower pre-defined limits are required in these statistical analyses.

2.3.5 Report Results

Shlomo Sasson: All reports of the results by the three laboratories are extensive and provide a solid basis for intra- and interlaboratory analyses of the data. Attention should be given to the

suggestions related to **improving initial data manipulations** and **curve fitting** (Appendix 5/pp. 16-17).

2.3.6 Provide any additional advice regarding the protocol.

Patrick Balaguer: ASSAY protocol

Conduct abbreviated pilot studies

As suggested in RTI project (appendix 5- overall report, 2n), a first screening of all compounds at high concentrations (1 mM and 0,1 mM for example) would enable to identify binders and indicate which compounds would have benefited from adjustement to lower concentrations. This should significantly cut down on the number of reruns required and the cost of characterizing compounds.

Reduction of the number of points by run.

In the analyses produced by the individual laboratories in this study, the top plateaus for the standard chemicals often exceeded the performance criteria by several tens of percentage points. The reason for the high plateaus may be related to the solvent control tubes. Such tubes placed at the end of the run (of several test chemicals run simultaneously) often yielded lower dpms than similar tubes placed at the beginning of the run. The average of all solvent control tubes was therefore lower than it would have been had only the first solvent control tubes been included. The lower average could have contributed to the appearance of higher-than-solvent-control values for the estradiol and norethynodrel. Processing this large number of tubes may have increased variability due to such factors as increased duration of exposure to room temperature (and subsequent denaturation of the receptor), and diminished ability to monitor partial pellet loss after centrifugation. This potential source of variability is expected to be less of a factor for laboratories if only one chemical is being tested at a time.

Consider not running all the standards with every assay.

Limit to running E2 and NOR once with each new cytosolic preparation and only E2 with each assay would reduce the number of points by run and would ameliorate the results.

Variability due to small amount (10 microliters) of test chemical and standard.

Choice of an higher volume (50 microliters) would reduce variability.

Radioactive estradiol concentration could be decreased.

The use of radioactive estradiol (^{3}H-E2) 1nM in the assay is perhaps too high to identify weak binders. Due to solubility, tested compounds cannot use at higher concentrations than 1 mM. At this concentration, weak binders can be classified as unequivocal (due to lower than 50% displacement). Use of radioactive estradiol 0,1 nM would increase the sensibility of the assay and enable to better characterize weak binders. However use of 0,1 nM ^{3}H-E2 will give lower signal than with 1 nM estradiol and an higher depletion by the receptor.

Remark 2. Use of estrogen receptors from another source than rat uterus.

Rat uterine cytosol contains both ER alpha and beta. These two receptors have not the same affinity for some natural, industrial or pharmaceuticals compounds like phytoestrogens, biphenols, cosmetics and ethynylestradiol (Kuiper et al, 1995, Paris et al, 2001, Escande et al, 2006, Molina et al, 2008). Variability of relative binding affinities (RBA) results can be induced by variation of the ER alpha /ER beta ratio in the different rat uterus cytosol batches. Thus, the protocol could be improved by using recombinant human ER alpha (or ER beta) receptor (Gangloff et al, 2001, Eiler et al, 2001, Pillon et al, 2005). Recombinant ER is tagged with six histines (6-His) or Glutatione-S-transferase (GST), which enables purification by Nickel- or glutathion-Sepharose. Advantages of purified ER are the producion of reproductible, well charaterized batches and a reduced variability of the assay by limiting binding interference with non ER proteins.

Alternatively, human cell lines (HeLa, U2OS) expressing ER alpha or ER beta (Escande at al, Sotoca et al, 2008) can also be used as source of estrogen receptor for binding experiments. In these cells, binding experiments could be done in lysed cells or whole-cells (Escande et al, 2006, Molina et al, 2008).

Advantages to use whole-cells is a simpler protocol with higher high-throughput screening possibilites. Inconvenients are concentrations of compounds that cannnot exceed 10^{-5}-10^{-4}M when lived cells are used.

Shlomo Sasson: 1. The choice of ovariectomized Sprague-Dawley rats for uteri collection is good. However, the argument made by the RTI scientists to extend the 8-day period after ovariectomy before excision of uteri due to remaining endogenous estrogens (Appendix 5/p.5) is not pharmacokinitically sound: the half-life of 17β-estradiol in Sprague-Dawley rats is about 10 hours (*Petroff and Mizing. Reproductive Biology, 3:131, 2003*). Thus, a full clearance of 17β-estradiol is expected after nearly 20 half-lives within the 8 days following ovariectomy. It is more likely, however, that the presence of stimulated uteri (due to fluid retention) was the consequence of an incomplete ovariectomy. In fact, RTI scientists confirmed the presence of residual ovarian tissue in some operated rats (*ibid*). It is therefore suggested that in cases where such remnants of ovarian tissue are found or when the excised uteri appear imbibed to discard them from the assay.

2. The requirement to perform a saturation binding assay for each RUC preparation prior to the competitive binding assay is mandatory: it allows the determination of the ER concentration in cytosols in term of maximal binding capacity of [^{3}H]-17β-estradiol. The large variation in the total protein content of the various rat uterine cytosols prepared by the 3 participating laboratories, particularly in the three cytosols prepared in Lab Y (ISR, Table 22, p.53) attests to this need. Despite the extraordinary high protein content in 2 of the cytosols in Lab Y, the mean maximal binding (B*max*) reported was 2-3-fold lower that that reported by Lab X, whose cytosols contained significantly lower protein content. This is probably due to the addition of a fixed amount of cytosolic proteins to the test tubes, leading inevitably to low content of ER. Therefore, I suggest to pre-determine the range of acceptable protein content in rat uterine cytosols. Large variations in protein content as reported by participating laboratories may complicate the interpretation of results.

3. It is important to keep the total receptor concentration in the binding assay low enough to restrict the binding interaction of 17β-estradiol with the receptor to a simple (single and non-interacting binding sites) mechanism (linear Scatchard plot, Hill coefficient=1) and to avoid complex binding kinetics due to receptor dimerization that occur at higher receptor concentration. It is therefore recommended to scrutinize all experiments where the Hill coefficient of 17β-estradiol binding is significantly higher than 1.0 and dilute the cytosol

accordingly. Dilution of the ER in the competitive binding assay to 0.5 nM usually suffices the requirement of a simple binding kinetics.

4. Often, the 'top plateau levels' of competition curves were markedly higher than the expected 100%. There are several technical explanations to these results, such as, underestimation of the control binding, receptor instability or variable concentration of solvents in test tubes. The recommendation to use a fixed volume of solvent (2 % ethanol) in all tubes, including controls, will eliminate the latter cause. It is suggested to design and include 'receptor stability' controls in all assays. The simplest assay requires an incubation of the cytosol under the same assay conditions of the competitive binding assay (16-20 hours at 4ºC) with no ligands. This cytosol is then used in a binding assay along with a freshly thawed cytosol. The binding capacity of these two cytosols is compared in a binding assay with saturating concentrations of [^{3}H]-17β-estradiol, an efficient anti-protease cocktail and incubation at high temperature (15-20ºC) for 1-2 hours followed by the HAP treatment. This simple assay eliminates misinterpretation of data due to receptor instability during the long incubation period at 4ºC.

5. By definition, if the [^{3}H]-17β-estradiol is present at a concentration that saturates the receptor- an equal concentration of unlabeled 17β-estradiol is expected to reduce the specific binding of [^{3}H]-17β-estradiol by 50% due to the 1:1 dilution of the radioactive ligand. In the case of the presence of excessive binding sites (non-saturating conditions) the 50% competition of the labeled 17β-estradiol with the unlabelled 17β-estratiol is observed at a ratio higher than 1:1 for the unlabeled estradiol. This may lead to an inaccurate estimation of IC_{50} values. The range of IC_{50} values is depicted in the Acceptable Standard 17β-estradiol Curves (Appendix 13/pp.1-3): the range reported by TRI is narrow enough and agrees with the assay's terms. However, the ranges shown by TRL and Hammer cover an entire order of magnitude. These wide ranges most probably result from variable receptor levels and non-saturating [^{3}H]-17β-estradiol concentrations. Therefore, I suggest using the data obtained from the standard 17β-estradiol saturation curves to calculate and use the optimal receptor concentration for maximal binding capacity (saturation) of [^{3}H]-17β-estradiol. In Appendix 6 (Appendix D, p.3, 8 and 12) Lab X presents ER binding curves that were saturates with 2-3 nM 17β-estradiol. Lab Y presents a summary of 3 saturation curves in Appendix 7 (p.16) in which maximal binding was observed with 0.5-1.0 nM [^{3}H]-17β-estradiol. Lab Z gives 4 saturation curves (Appendix 8/p. A-17, C-20,

D-17, E-17) which did not reach saturation at 1-2 nM [^{3}H]-17β-estradiol and one curve (B-22) that saturated between 1.3-3.6 nM [^{3}H]-17β-estradiol. These inconsistent data may explain intra- and interlaboratory variations of the calculated kinetic parameters (K*d*, IC_{50}, and RBA). The equivocal results of the competitive binding assay of Test Chemical #1 (17β-estradiol) presented by Lab Z could have been avoided if the assays were conducted with a prescribed concentration ER and saturating concentrations [^{3}H]-17β-estradiol.

6. It is highly recommended that researchers will assess potential precipitation of hydrophobic test compounds in the binding assay mixture by using physical methods such as light scattering spectroscopy. Precipitation problems can often be solved with some organic solvents. Appendix 3 shows the compatibility of the binding assay to 2% ethanol. However, the recommendation to allow DMSO up to 20% in the assay is worrying; the U-shaped displacement curves of various ligands to the ER (Figures16, 18, 22 and 23) indicate that DMSO may affect the receptor and change its binding kinetic properties in a concentration-dependent manner. Indeed, such effects of solvents (*i.e*., dimethylformamide) on the binding kinetics of the ER were previously reported (*Sasson & Notides, J. Steroid Biochem. 29:491-5, 1988*).

7. The absorption of hydrophobic chemicals to test tube walls is not always eliminated by using siliconized glass (borosilicate) tubes. Compounds with a high partition coefficient (*e.g.,* the Log*P* of tamoxifen is 6.58 in comparison with 3.67 of 17β-estradiol) do interact with these seemingly 'inert' surfaces. This absorption problem may be solved by other means than just using solvents at different concentration. It was shown (*ibid*) that a protein content higher than 2.5 mg/ml in the binding assay reduced significantly tamoxifen lose to the tube walls. Various proteins, such as insulin or IgG, which minimally increase the non-specific binding of 17β-estradiol, can be added to low-protein cytosols for this purpose. Yet, such unusual solutions are rare because ethanol is a good solvent for most chemicals and is well tolerated in the binding assay up to a final volume of 2%.

8. The HAP sedimentation procedure is used for separating free from bound [^{3}H]-17β-estradiol. Following washes the HAP slurry is extracted with ethanol and a 1 ml aliquot is then mixed with 14 ml of scintillation cocktail to measure radioactivity in a β-counter. An accurate conversion of counts-per-minute (cpm) to disintegrations-per-minute (dpm) is of a paramount importance for subsequent data analysis. The counting efficiency may differ among various scintillation

cocktails and β-counter settings. Various methods are available to estimate counting efficiency and correctly calculate dpm: internal standards, Channel Ratio (CR) or external standard (γ-radiation). It is imperative that each participating laboratory be able to reliably calculate dpm of radioactive samples. The use of internal standards (that is, fixed amount of μCi of [^{3}H]-17β-estradiol counted with 1 ml ethanol in 14 ml of scintillation cocktail) is often adequate; Erratic CR values identify counting vials in which quenching was irregular.

9. RTI scientists listed several suggestions regarding the assay protocol (Appendix 5/p. 15). The suggestion to **conduct an abbreviated pilot study** (screen test compounds at 10 μM and further characterize only those that inhibit [^{3}H]-17β-estradiol binding by 50% or more) sounds reasonable. However, the limit in my opinion should be lower than 50% to eliminate false negative determinations from the assay. The suggestion to **determine the K*d* values of [^{3}H]-17β-estradiol only once** for each newly prepared cytosol is acceptable, as long as the B*max* values do not significantly change when the cytosol is tested repeatedly in independent assays. Moreover, inclusion of a control for receptor stability (see above) will also warrant similar experimental conditions among assays. The suggestion **not to run full standard assays** (17β-estradiol and norethynodrel), but use the compounds at their respective IC_{20} and IC_{80} values is problematic: close examinations of the acceptable standard curves in various runs of each participating laboratory and among them fails to provide absolute concentrations of each compound that induce 20 or 80% displacement of [^{3}H]-17β-estradiol binding. Furthermore, full standard curves and the experimental values of IC_{50} of unlabelled 17β-estradiol do provide important information on the quality and saturability of the assay. Similarly, a Hill coefficient value close to 1 indicates a simple binding mechanism. I do, however, tend to agree that the norethynodrel standard curve is redundant. **Stability of materials** is very important and I fully support the suggestions regarding this issue. Similarly, problems with **solubility of chemicals** are critical and the suggestion to employ various approaches to solubilize the compounds, other than these specified in the protocols, is sound. Yet, this requires control experiments showing that the solubilizing solvent or other compounds used do not interfere with the binding kinetics of 17β-estradiol. Obviously, **brands of reagents and suppliers** do vary. Thus, the idea to provide a "suggested supply list" is encouraged. The problem regarding **difficulties in pipetting 10 μL of test chemical and standards** can easily be solved if these test chemicals or standards are pipetted first to dry tube- no immersion of pipette tips in the binding mixture or an

incomplete delivery of the tip's content are encountered this way. Accordingly, the order of addition of the assay components should be modified accordingly.

10. The case of dual (mixed) estrogenic-antiestrogenic properties of certain ER ligands was not addressed in the study. Some compounds, like estriol, act as weak estrogens when present or administered alone to immature or ovariectomized animals. However, when present in excess over estradiol, these compounds exert potent antiestrogenic properties. This duality in function depends on the concentration of the competing ligand relative to the ambient 17β-estradiol concentration. This phenomenon results from a partial activation of the ER when the test compound is present alone, and to an aberrant dimrization of the ER in the presence of the natural ligand (*Melamed et al., Mol. Endocrinol. 11:1868-78, 1997*). The simple binding assay employed in this study is not aimed at analyzing such complex kinetic interactions. However, these complex interactions may occur in animal studies. Hence, it is recommended that functional assays aimed at determining the properties of ER binders *in vitro* and *in vivo* would carefully explore their potential to act as mixed agonists/antagonists. Equally important is the potential of such compounds to exert antiestrogenic properties in female with normal ovary function and secretion of 17β-estradiol, while acting as estrogenic agonists in males, due to the absence of significant levels of natural estrogens. This spectrum of functions should be acknowledged and examined carefully in studies that aim at determining biological and toxicological functions of such ER ligands.

Marie-Louise Scippo: *1. Data for plotting the competitive binding curves : the y axis.*
In Appendix 1, page 46, paragraph 10.7, the data analysis is not totally clearly explained.

It is indicated that "*The competitive binding curve is plotted as specific [^{3}H]-17β-estradiol binding versus the concentration (log10 units) of the competitor*."
If we look at 10.7.1. "*Terminology*", the definition of the specific binding is "*Total binding minus non-specific binding*", but the "*total binding*" is not defined, and thus the way to calculate the specific binding is not indicated.

It should be added that :

Total binding is radioactivity in DPMs in the tube that contains [3H]-17β-estradiol and receptor, in the absence or presence of competitor.

Furthermore, it is not clear which figures we have to plot on the y axis of the competitive binding curve.

In Appendix 4, page 14, Paragraph 2.1.1. (Input data specification), it is indicated that "*input data for the dependent variables should be standardized and expressed as "% binding of the reference ligand to the receptor*", which I translate as the ratio :

$$\frac{\text{Total binding (in presence of competitor) - NSB}}{\text{Total binding (in absence of competitor) - NSB}}$$

As this is indicated nowhere in the documents (may be it is clear on the excel worksheets, but unfortunately, we didn't receive the excel files, this should be added to the terminology, under "*specific binding*".

2. Receptor concentration in the assay

2.1. Performance criteria of the saturation binding assay: Standardization of the receptor concentration.

In Appendix 1, page 20, 9.1.5. "*Standardization of receptor concentration*", it is indicated that "*For the saturation assay, the optimal protein concentration binds 25 -35% of the total radiolabeled estradiol that has been added to the tube. To ensure that this percent range of radioligand is bound at the lowest concentration of radioligand added to the assay, the 0.03 nM concentration shall be used to make this determination for the saturation binding assay*".

Under 9.7. ("*test report*", pages 26 - 27 of Appendix 1), it is not clearly asked to the lab to report the percent of bound radioligand at the 0.03 M concentration.
It should be clearly asked to the labs to report this value.

Furthermore, I tried to retrieve these data in the raw data of the individual labs, and I found for both lab Z and Y (Appendix 8, page A-2 and Appendix 7, page B-9 respectively), that the percent of radioligand bound at 0.03 M radioinert estradiol was indicated in a column named "*10*

Percent rule", and the criteria was obviously that the percent of radioligand bound should be below 10%. This is not in agreement with the protocol (which requires for 25 - 35 %).

2.2. Performance criteria of the saturation binding assay: Standardization of the receptor concentration.

In Appendix 1, Page 47, 10.7.3. "*Performance criteria for the competitive binding assay*", it is indicated that: "*Ligand depletion is minimal. Specifically, the ratio of total binding in the absence of competitor to the total amount of [^{3}H]-17β-estradiol added per assay tube is no greater than approximately 15%*".

Again here, it is not asked to the labs to report this ratio, but only the amount of protein added per tube.

It should be asked to the lab to report this parameter too.

Again here, lab Y used the below "*10% rule*", and generally, the percent of radioligand bound was around 3%, which I find too low (Appendix 7, page D-3).

I was not able to find the information in the report of lab X.

Furthermore, I see here a contradiction with Appendix 1, page 20, 9.1.5. where it is said that "*For the saturation assay, the optimal protein concentration binds 25 -35% of the total radiolabeled estradiol that has been added to the tube*". This would mean that a different protein concentration should be used for saturation binding assays (25 - 35% binding of the total radioligand added) and for competitive binding assays (15% binding of the total radioligand added), which is not what is done in practice. Please, clarify that point in the protocol.

3. In a validation study, the protocol given to the participating laboratories should be a definitive one. As several changes have been made in the protocol after the second interlaboratory study (we see that by comparing the modified protocol described in Appendix 1 and the original protocol described in Appendix 5), it should be relevant to organize a third interlaboratory study using (and reviewed) a definitive protocol.

A specific remark :

Appendix 1, page 48, table 9 : the lower limit of the top plateau for estradiol is 90% and not 94%.

2.4 Comment on Whether the Strengths and/or Limitations of the Assay Have Been Adequately Addressed

Patrick Balaguer: Strenghts.

As an in vitro assay, the ER-RUC provides direct interaction between chemical and ER. The assay provides consistent responses at the simple screening level, across laboratories, and these responses are in line with expectations for those chemicals tested whose ER binding behavior is well-established. Compounds that do not interact with ER consistently test negative in the assay. The assay is short and inexpensive compared to in vivo tests.

Limitations.

The assay is sensitive to many details of preparation and technique and can show wide variability if not performed exactly as stated in the protocol. It is subjected to problems if the receptor concentration in the cytosol is too low or too high. Another problem is that the tubes are not kept cold at all times during preparation, incubation and separation of bound from free tracer. This problem is certainly responsible of the the top plateaus for the standart chemicals (estradiol and norethynodrel) often exceeded the performance criteria by several tens of percentage points. However, the data suggest that a lab that meets the performance criteria for teh standart and weak positive is likely to generate data that is much less varaiable than laboratories that do not meet performance criteria.

The analysis of datasets was relatively complicated. The standardization of the assay does not allow precise and replicable quantitative analysis of log(IC50) and RBA.

Finally, the assay requires the use of animals.

Ingemar Pongratz: Section 2

I feel that the assay as first line approach is valid, but would benefit from additional scientific experimentation and validation, in particular characterize some of the issues regarding the circadian expression of the ERβ isofom and of the potential stability of the hsp90-ERα and

hsp90 ERβ complex. I feel this limitations should be addressed in the text and efforts should be undertaken to fill the scientific gap regarding these issues.

A further limitation of the assay is due to the use of laboratory animals. There are a number of parameters that are difficult to control, such as quality of feed. Has the possibility to use established cell-lines been considered?

Shlomo Sasson: The strengths of the assay are adequately addressed (ISR, pp.68-9). The major strength was the ability of the RUC-ER assay to classify correctly most test compounds, despite inconsistencies among the participating laboratories and various technical problems. Clearly, once the recombinant hER binding assay is completed, a thorough comparison and analysis of the results of the two independent assays is required for further validation of the RUC-ER as a relatively simple and affordable screening assay. The weaknesses of the assay (*ibid*, p.69-70) refer mostly to technical and methodological aspects such as, ER concentration in the assay, insolubility of test compounds, complicated analyses of data and the inevitable use of rats. There are no satisfactory explanations to the lack of adherence of some participating laboratories to the standard protocols.

Marie-Louise Scippo: The strengths and the weaknesses have been adequately addressed. The strengths of the assay are, for the most, relevant for all estrogen receptor binding assays but one important specificity of the ER-RUC is stressed: it is that the rat uterine cytosol contains both isoforms alpha and beta of the estrogen receptor. The high degree of homology between the rat and the human ER ligand binding domain could also be mentioned here.

The main weakness of the assay comes from the use of animals to prepare the binding fraction, with not only ethical, but also technical consequences (lack of reproducibility in receptor preparations). The solution is the use of recombinant ER binding assays, using both alpha and beta isoforms of the ER.

If a third interlaboratory study is organized, it would be interesting to analyze the same list of chemicals with both ER-RUC and recombinant ER binding assays. The comparison could provide different conclusions that could add some weight to the strength of the ER-RUC using

both α and β isoforms of ER compared to a recombinant ER binding assay using the α isoform only.

William J. Welsh: The strengths and limitations of the assay are not clearly specified the ISR. This should be done.

2.5 Provide Comments on the Impacts of the Choice of a) Test Substances, b) Analytical Methods, and c)Statistical Methods in Terms of Demonstrating the Performance of the Assay

Patrick Balaguer: Yes, the test substances were appropriately chosen to demonstrate the performance of the assay. The test substances correspond to a wide range of strenghs and chemical structure. They are 3 very strong, 3 strong (including E2), 3 moderate, 9 weak and 5 negative compounds. Among negative compounds, atrazine (non-binder but estrogen-active) was included.
Analytical methods and statistical methods were also appropriately chosen to demonstrate the performance of the assay.

Ingemar Pongratz: I think that both the test substances and the methods are appropiate for the purpose.

I have little experience with the statiscal methodology so have no opinion regarding the statistical methods that were chosen for the assay.

Shlomo Sasson: Test substances were carefully collected to cover a wide range of compounds that bind to the ER at very strong (higher than 17β-estradiol), moderate or weak affinities, regardless of their biological functions. The analytical methods are appropriate to demonstrate the performance of the assay. As stated above, I refrain from commenting on the statistical methods.

Marie-Louise Scippo: The choice of a panel of strong, very strong, moderate, weak or negative binders seems relevant.

However, it would have been better to avoid substances without available data in the literature about their binding capacity (in ER-RUC assays) to the ER.

In ISR, page 52, table 21, page 52, references of "*historical*" RBAs should be given.

William J. Welsh: Yes, these three aspects have been amply presented in the ISR. The substances and methods are appropriate to demonstrate the performance of the assay.

2.6 Provide Comments on Repeatability and Reproducibility of the Results Obtained with the Assay, Considering the Variability Inherent in the Biological and Chemical Test Methods

Patrick Balaguer: Considering the variability inherent in biological and chemical test methods, the results obtained with this assay are relatively repeatable and reproducible.

Very strong an strong compounds.

All of very strong, strong and moderate (17b-estradiol, 17-ethynylestradiol, DES, Meso-hexestrol, zearalenone, tamoxifen, norethynodrel, genistein and equol) binders were correctly determined to interact with rat uterine ER. The only exception is estradiol in laboratory Z. Binding at the concentrations tested (100 pM to 1 mM) showed clear interaction with the receptor but the laboratory did not adjust the test concentration range to a more dilute range that would have allowed characterization of the full binding curve as required by the protocol.

Note that laboratory Z misinterpretated (compare table 27 ER-Ruc and table 2 p22 appendix 8) strong affinity compound 1 (estradiol), 3 (DES) and 13 (tamoxifen). Note that tamoxifen is not by itself a strong binder. 4-hydroxy tamoxifen (a tamoxifen metabolite) is the strong binder.

When the three laboratories' RBAs were compared, similar values were obtained for very strong and strong compounds.

Moderate compounds

For moderate compounds, similar results were obtained for norethynodrel, equol and zearalenone while variability between laboratories is observed for genistein and tamoxifen (appendix 6 table 5, appendix 7 table 6 and appendix 8 table 2).

Weak binders

The weak chemicals had various classifications by the laboratories. Butyl paraben and bisphenol A were classified as positive by all of the laboratories.

Nonyphenol (mixture), 4-n-heptylphenol, and enterolactone were positive, negative and equivocal for laboratory X, Y, and Z, respectively. Kepone was positive in two laboratories and equivocal in the third. O,p'-DDT was negative in two laboratories and positive in the third. After analysis of the laboratory results, it was concluded in the ER-RUC ISR that these seven compounds were binders. I agree with the ER-RIC ISR exepted for o,p'DDT for which more experiments are necessary to conclude.

5alpha-dihydrotestosterone was negative in two laboratories and equivocal in the third. Benz(a)anthracene was classified as negative by all of the laboratories.

Thus, most of the compounds that were expected to be weak binder were «positive« in their responses. However as mentionned above, the assay does not allow precise and replicable quantitative analysis of log(IC50) and RBA.

For the negative chemicals (corticosterone, progesterone, octyltriethoxysilane, atrazine and R1881), laboratory X and Y had corticosterone listed as positive, and laboratory Z had it as equivocal. Curiously, this compound was correctly classified as negative in the optional portion.

Progesterone and octyltriethoxysilane were classified as negative by all three laboratories. Atrazine was classified correctly as negative by laboratory T and Z but positive by laboratory X.

R1881 that was designed to be a negative control for the assay proved to be positive for the control runs, and as a coded chemical it was positive for laboratory X and Y and equivocal for laboratory Z.

Thus, most of the compounds that were expected not to be binder were « negative» or « equivocal » in their responses.

Ingemar Pongratz: In general I would say yes. I would suggest that the experimental issues raised previously be taken into account.

The possible effect of the circadian system can be covered by characterizeng the perak of ERβ expression at different time points, the food composition of the animals should be centralized, and stabiltiy of the hsp90 complex needs to be better addressed. A experimental possibility would be to use cell-lines, I understand however that this would lead to less material being available.

Shlomo Sasson: In most cases the answer to this question is positive. The summary Table (Table 27, ISR, p.60) shows that despite a considerable variability in the data obtained by the 3 laboratories, most test compounds were correctly classified. The case of Compound #1 (17β-estradiol) in Lab Z that was classified 'equivocal' is peculiar and most probably results from deviations from the standard protocol of the competitive binding assay. In fact, it is somewhat surprising that despite considerable variability in datasets obtain by participating laboratories and among them- the final evaluation was powerful enough to categorize most test chemicals according to their rank. Yet, adherence to the protocols and further optimization of the assay will eventually lead to more reproducible results.

Marie-Louise Scippo: The results were sufficiently repeatable and reproducible for the reference compounds (estradiol and norethynodrel), but for test chemicals, some results were too much dispersed (for example, for test chemical n°1, for lab Z, see in Appendix 5, page 1 of Appendix 2).

William J. Welsh: Yes, the results from the assay were extremely repeatable and reproducible especially in view of the variability inherent in biological and chemical test methods.

2.7 Comment on Whether the Appropriate Parameters were Selected and Reasonable Values Chosen to Ensure Proper Performance of the Assay, with Respect to the Performance Criteria

Patrick Balaguer: Yes, appropriate parameters were selected and reasonable values were chosen to ensure proper performance of the assay.

Ingemar Pongratz: In my opinion yes.

Shlomo Sasson: The saturation binding assay of 17β-estradiol and the competitive binding assay were introduced some 40 years ago with the discovery of the estrogen receptor and the availability of radioactive estradiol. The kinetic parameters, K*d*, B*max* and RBA have been standard endpoints in numerous binding assays. The K*d* values for 17β-estradiol binding to the RUC-ER in this study fall within the acceptable range. Obviously, B*max* values can vary widely, depending on the quality and ER concentration of the uterine cytosols.

Marie-Louise Scippo: Performance criteria were relevant, but I would add some criteria such as those presented in the answer to question n°3.

To ensure the performance of the assay to detect ER interactive chemicals, I would add some criteria such as maximum rate of false negative and false positive decisions, measured with known substances.

In this validation study, 23 test chemicals were used to check this parameter, but the results are not very conclusive. On 22 chemicals (if we exclude R1881), respectively 5, 2 and 5 false decisions were taken by labs X, Y and Z (see ISR, page 60, table 27).

This rate of false decision will also depend of the weight of the ER-RUC assay in the Tier 1 screening (see my answer to question 1).

William J. Welsh: Yes, appropriate performance criteria were selected to ensure proper performance of the assay. These criteria are clearly articulated throughout the ISR.

2.8 Comment on the Clarity, Comprehensiveness and Consistency of the Data Interpretation with the Stated Purpose of the Assay

Patrick Balaguer: Yes, the data interpretation criteria are clear, comprehensive, and consistent with the stated purpose.

Ingemar Pongratz: The data interpretation is in my opinion clear and consistent with the purpose.

Shlomo Sasson: The data interpretation follows rigorous kinetic and statistical analyses. Both provide clear representation and analyses of the data and are fully consistent with the stated purpose of the assay.

Marie-Louise Scippo: Yes, but there are some inconsistencies between the different documents.

Appendix 5 presents the report of the second inter-laboratory validation of the estrogen receptor binding assay (rat uterin cytosol). First, in Appendix 5, labs are named A, B, C, which is confusing for the reader, because in the ISR, the labs are named X, Z and Y respectively.

In Appendix 5, page 10, it is indicated that the labs were required to classify the unknown chemicals as positive, negative or equivocal binders. This is not the classification described in the protocol (Appendix 1, page 48,10.7.4), where the classification is described as interactive, not interactive or equivocal.

The results of the classification of the 23 test chemicals by the 3 labs X, Y and Z are presented in both ISR and Appendix 5 documents, in table 27 and 5 respectively. I expected that the results were the same in both documents, but they are not for 14 chemicals out of 23, without any explanation in the ISR. This is very confusing for the reader!

William J. Welsh: Yes, the criteria for data interpretation are described in a clear, comprehensive and consistent manner.

2.9 Please Comment on the Overall Utility of the Assay as a Screening Tool, to be used by the EPA, to Identify Chemicals that have the Potential to Interact with the Endocrine System Sufficiently to Warrant Further Testing

Patrick Balaguer: The estrogen receptor binding assay using rat uterine cytosol will represent a validated assay for simple screening for interaction with the estrogen receptor in the context of a battery of in vitro assays. EPA has optimized and standardized the most important parameters of this assay and has shown that the resulting protocol is transferable to others laboratories and enable to identify chemicals that have the potential to interact with the endocrine system.

Analysis of compounds binding curves in appendix II of Appendix 5-overall report, enable easily to identify the nature of the compounds.
Among very strong and strong binders (17b-estradiol, 17-ethynylestradiol, DES, meso-hexestrol, zearalenone and tamoxifen), zearalenone and tamoxifen are not strong binders (IC50s around 100 nM).

Among moderate binders (genistein, norethynodrel, equol), genistein (IC50 around 100 nM) is a stronger binder than norethynodrel an equol (IC50s around 1000 nM).

Among predicted weak binders (butyl paraben, nonylphenol, o,p'-DDT, 5alpha-dihdrosterone, bisphenol A, 4-n-heptylphenol, kepone, benz(a)thracene, enterolactone), it was very easy to identify real weak binders in the test (butyl paraben, nonylphenol, bisphenol A, 4-n-heptylphenol, kepone, enterolactone) and non binders (o,p'-DDT, 5alpha-dihydrosterone, benz(a)thracene).

Finally among predicted negative compounds (atrazine, corticosterone, octyltriethoxysilane, progesterone and R1881), it was easy to identify weak binders (corticosterone and R1881) and real negative compounds (atrazine, octyltriethoxysilane and progesterone). Curiously, corticosterone was negative when the tested optional chemicals. Concerning R1881, it is likely that the expectation of non-binding is not correct.

When the analysis was developed, the EPA expectation was that standardization of the assay would allow precise and replicable quantitative analysis of log(IC50)s and Relative Binding Affinities. The expectation was also that precise, standardized methods of analysis would contribute to reproducibility and therefore use in other applications such as structure-activity relationship models.

The variability of the results of the assay may not support for the moment use for quantitative structure-activity relationship model development. Even if strenght of binding could be evaluated, high variability of compounds RBAs were observed. In the face of the variability encountered, EPA is assessing whether such analysis could be replaced with a simpler analysis and still meet the needs of the Screening Program.

However, it should be remembered that while intralaboratory variability was disappointingly high for at least one laboratory in this study, such variability is not expected to be as much of a problem for laboratories that demonstrate the ability to meet the required performance criteria. The limited time available to run this large study on 23 chemicals apparently did not allow development of the proficiency necessary to obtain precise runs in all laboratories. The fact that results were almost all in accord with expectations when screening for interaction despite the variability in quantitative values shows that the assay is robust for this use.

Furthermore, use of recombinant receptor rather than receptor obtained from whole animals will enable to use purified receptor which would decrease the variability of the binding assay.
Finally, small ameliorations of the protocol, fine adjustement of performance criteria and increase of experience by the laboratories should allow the test to reach the expected purpose which is to determine the potential of a substance to interact with the endocrine system.

Ingemar Pongratz: I think the assay is suitable to identify compounds that bind to the estrogen receptors ERα and ERβ. Again, the assay should be regarded in the context of being one of several assays. I have however certain reservations to the statement that the assay will cover metabolism. Metabolism and the presence of the p450 enzymes is in many cases a cell and tissue specific process and all the necessary components may not be present in uterine tissue and cells.

Shlomo Sasson: This assay is one in a battery of assays aimed at identifying endocrine disruptors. The RUC-ER binding assay, which was developed some 40 years ago, provides simple kinetic analyses of ER binding interactions with natural and synthetic agonists and antagonists. The principles of the assay have not significantly changed over this period. A major development is the source of the ER: uteri from various species (mostly rats and calves) have been used extensively for years. The current availability of recombinant human ERα and ERβ now enables studies more relevant to human. Thus, it will be interesting to compare the results of the RUC-ER assay to the recombinant hER project that is run in parallel. Both assays are relatively simple and can be included in the battery of assays to provide a first-line screening of potential ligands to the ER. I believe that the RUC-ER meets these criteria. Nevertheless, the protocols of the assay require some further optimization.

Marie-Louise Scippo: The introduction clearly indicates that the ER-RUC assay is one of the complementary assays included in Tier-1, listed in table 1, page 3. In page 1, it is indicated that "*A negative result in Tier 1 would be sufficient to put a chemical aside as having low to no potential to cause endocrine disruption, whereas a positive result would require further testing in Tier 2* ". And in page 7, we read : "*An individual assay may serve to strengthen the weight of evidence in a determination (e.g., positive results in an ER binding assay in conjunction with positive results in the uterotropic and pubertal female assays would provide a consistent signal for estrogenicity)".* However, it is not clear, how is taken the final decision after Tier-1 (negative result or positive result), if all the complementary assays of the Tier-1 battery don't give concordant results (for example positive results in the ER binding assay but negative result in the in vitro transcriptional assay).

In page 2 of the ISR, there is a mention about "*false negatives*" and "*false positives*" resulting from the Tier-1 screening in the sentence "*Maximum sensitivity to minimize false negatives while permitting an as yet undetermined, but acceptable, level of false positives*". It would be necessary to explain how a positive or a negative decision is taken after the Tier-1 screening, or to mention a reference giving this explanation.

William J. Welsh: The ER binding assay is fairly adequate as an initial screening tool. However, the ISR should describe to what extent the results (and/or interpretation of the results) of the assay might be influenced by recognition of non-genomic estrogenic signaling effects.

2.10 Additional Comments and Materials Submitted

Marie-Louise Scippo: Additional remarks about the integrated summary report (file ER-RUC ISR, v3.16c.doc)

- Please, add a list of acronyms and abbreviations, it would be easier for the reader.
- Page 7, the reference "*USEPA, 2008*" is indicated in blue and underlined, but there is no hyperlink.
- Page 9, last paragraph of "*V. Overview of the assay*": In the sentence: "*The competitive assay measures the binding of [^{3}H]-estradiol at a fixed concentration in the presence of a wide range (eight orders of magnitude) of test chemical concentration*". Eight orders of magnitude should be replaced by seven orders of magnitude since the concentrations tested are from 100 pM to 1 mM (see section F, page 16, and Appendix 1).
- Page 10: in the sentence: "*The solvent used for a test chemical must also be used for the reference chemical (inert 17β-estradiol) and the control chemicals (norethynodrel and octyltriethoxysilane) unless the solvent is water*." replace "*inert 17β-estradiol*" by "radioinert 17β-estradiol".
- Same remark for the table in page 12.
- Page 15, in the sentence: "*If the test substance interacts with the receptor, it inhibits the binding of increasing amounts of radiolabeled estradiol*." please delete "*increasing amounts*", because it is a fixed amount of radiolabeled estradiol.
- Page 19, table 3: "*Performance criteria for competitive binding, reference and weak positive controls*". The figures of table 3 are different from figures of table 16 : "*Performance criteria for second interlaboratory study*". Is table 3 valid for the 1st interlaboratory study only? As I understood, octyltriethoxysilane was not used as a negative control during the 1st interlaboratory. And in table 16, R1881 is still mentioned as negative control. Please give explanations about the differences between table 3 and table 16. Furthermore, the upper limit for the top plateau level for norethynodrel (% binding) is 110 and not 10.

- Page 20, paragraph III.A. "*Buffer Composition and receptor concentration*". This paragraph presents the evaluation of the effect of the receptor concentration, during the optimization of the assay. It is concluded that the optimal concentration is 50 µg protein/tube. Nothing is clearly said about the importance of the percentage of binding of the radioligand. It would be clearer to indicate, as it is in the protocol (Appendix 1, page 20, paragraph 9.1.5.), that the optimal receptor concentration has to be adjusted to a quantity of protein that "*binds 25 -35% of the total radiolabeled estradiol that has been added to the tube*" (measured using 0.3 nM estradiol in saturation binding assays). Furthermore, the protocol (Appendix 1, page 25) indicates that the protein concentration corresponding to that binding is generally in the range between 35 and 100µg/tube. For competitive binding assays, the protocol (Appendix 1, page 33) indicates that "*the optimal amount of cytosolic protein added should contain enough receptor to bind no more than 10 - 15% of the radiolabeled estradiol that has been added to the tube*".
- Page 20: In the sentence (last sentence of the page): "*The original design called for 100 mg of protein per assay tube. Protein concentrations from 25 to 100 mg/tube were evaluated*.", please consider to replace "*mg*" per µg.
- Page 21: In the sentence : "*The dpm values of the 100% binding tubes in these assays ranged from 1 to 4 percent of those of the hot tubes, indicating that ligand depletion was not significant*.", please define what are the "*hot tubes*" and preferably don't use "*hot tubes*", which is laboratory "jargon".
- Pages 27 and 28, Figures 3 and 4 are not cited in the text.
- Page 45, under "*B. Second interlaboratory study*", in the sentence: "*Therefore in the second study one stock solution was made in solvent and this solution diluted sequentially with buffer*.", please, consider to replace "*buffer*" by "solvent".
- Page 46, table 16 is still mentioning R1881, even if it was finally not used as a negative control.
- Page 47, table 18, Lab X reported values for the "*Ki*", but nothing is said in the text about the Ki. It would be better to delete this row from the table, and if not, to give some comments about the Ki.
- Pages 48 and 49, table 19 and 20 show results of qualification runs of the 3 labs participating to the 2nd interlaboratory study. The results are compared with performance criteria described in table 16, and non compliant results are highlighted. In table 16, for

both estradiol and norethynodrel, lower and and upper limits for bottom plateau level are -5.0 and 1.0 % binding, respectively and lower and and upper limits for top plateau level are 90.0 and 110.0 % binding, respectively. In tables 19 and 20, cells containing figures not corresponding to these criteria are highlighted, but there are some inconsistencies (some cells should be highlighted and are not and others are highlighted and should not be), please check this.

3.0 PEER REVIEW COMMENTS ORGANIZED BY REVIEWER

Peer review comments received for the ER Binding assay are presented in the sub-sections below and are organized by reviewer. Peer review comments are presented in full, unedited text as received from each reviewer.

3.1 Patrick Balaguer Review Comments

1. Is the stated purpose of the assay clear?

Yes, the stated purpose of the assay is very clear. The purpose is to provide a test that enable to identify the potential of chemicals to interact with rat uterine estrogen receptors.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

Yes, the assay is biologically and toxicologically relevant to the stated purpose. It enables to identify binder, equivocal or non binders. However, it does not allow to quantitative structure-activity relationship development for the moment.

3. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:

a) comprehend the objective;

Yes, the description of the methodology is clear and concise and enable the comprehension of the objective.

However the finding that strong binders were considered « equivocal » rather than clearly interactive in one laboratory when tested blindy deserves and need to be noted in the protocol.

When binding at the concentrations tested showed clear interaction with the receptor, the laboratory did adjust the test concentration range to a more dilute range that would allow characterization of the full binding curve as required by the protocol. Thus, the protocol need to be adjusted to emphasize that a full curve must be obtained where there are clear indications of binding.

b) conduct the assay;

Yes, the description of the methodology is clear and concise and enable the conduction of the objective;

Again, the protocol need to be adjusted to emphasize that a full curve must be obtained where there are clear indications of binding.

c) observe and measure prescribed endpoints;

Yes, the description of the methodology is clear and concise and enable the observation and the measurement of prescribed endpoints;

Again, the protocol need to be adjusted to emphasize that if binding is clearly observed at 100 pM-1mM, the laboratory did adjust the test concentration range to a more dilute range.

d) compile and prepare data for statistical analyses; and report the results?

Yes, the description of the methodology is clear and concise and enable the compilation and the preparation of the data dor the statiscal analysis and the report of the results;

What additional advice, if any, can be given regarding the protocol?

<u>ASSAY protocol</u>

<u>Conduct abbreviated pilot studies</u>

As suggested in RTI project (appendix 5- overall report, 2n), a first screening of all compounds at high concentrations (1 mM and 0,1 mM for example) would enable to identify binders and indicate which compounds would have benefited from adjustement to lower concentrations. This should significantly cut down on the number of reruns required and the cost of characterizing compounds.

<u>Reduction of the number of points by run.</u>

In the analyses produced by the individual laboratories in this study, the top plateaus for the standard chemicals often exceeded the performance criteria by several tens of percentage points. The reason for the high plateaus may be related to the solvent control tubes. Such tubes placed at the end of the run (of several test chemicals run simultaneously) often yielded lower dpms than similar tubes placed

at the beginning of the run. The average of all solvent control tubes was therefore lower than it would have been had only the first solvent control tubes been included. The lower average could have contributed to the appearance of higher-than-solvent-control values for the estradiol and norethynodrel. Processing this large number of tubes may have increased variability due to such factors as increased duration of exposure to room temperature (and subsequent denaturation of the receptor), and diminished ability to monitor partial pellet loss after centrifugation. This potential source of variability is expected to be less of a factor for laboratories if only one chemical is being tested at a time.

Consider not running all the standards with every assay.

Limit to running E2 and NOR once with each new cytosolic preparation and only E2 with each assay would reduce the number of points by run and would ameliorate the results..

Variability due to small amount (10 microliters) of test chemical and standard.

Choice of an higher volume (50 microliters) would reduce variability.

Radioactive estradiol concentration could be decreased.

The use of radioactive estradiol (^{3}H-E2) 1nM in the assay is perhaps too high to identify weak binders. Due to solubility, tested compounds cannot use at higher concentrations than 1 mM. At this concentration, weak binders can be classified as unequivocal (due to lower than 50% displacement). Use of radioactive estradiol 0,1 nM would increase the sensibility of the assay and enable to better characterize weak binders. However use of 0,1 nM ^{3}H-E2 will give lower signal than with 1 nM estradiol and an higher depletion by the receptor.

Remark 2. Use of estrogen receptors from another source than rat uterus.

Rat uterine cytosol contains both ER alpha and beta. These two receptors have not the same affinity for some natural, industrial or pharmaceuticals compounds like phytoestrogens, biphenols, cosmetics and ethynylestradiol (Kuiper et al, 1995, Paris et al, 2001, Escande et al, 2006, Molina et al, 2008). Variability of relative binding affinities (RBA) results can be induced by variation of the ER alpha /ER beta ratio in the different rat uterus cytosol batches.

Thus, the protocol could be improved by using recombinant human ER alpha (or ER beta) receptor (Gangloff et al, 2001, Eiler et al, 2001, Pillon et al, 2005). Recombinant ER is tagged with six histines (6-His) or Glutatione-S-transferase (GST) which enables purification by Nickel- or glutathion-Sepharose. Advantages of purified ER are the producion of reproductible, well charaterized batches and a reduced variability of the assay by limiting binding interference with non ER proteins.

Alternatively, human cell lines (HeLa, U2OS) expressing ER alpha or ER beta (Escande at al, Sotoca et al, 2008) can also be used as source of estrogen receptor for binding experiments. In these cells, binding experiments could be done in lysed cells or whole-cells (Escande et al, 2006, Molina et al, 2008).

Advantages to use whole-cells is a simpler protocol with higher high-throughput screening possibilites. Inconvenients are concentrations of compounds that cannnot exceed 10^{-5}-10^{-4}M when lived cells are used.

4. Have the strengths and/or limitations of the assay been adequately addressed?

Strenghts.

As an in vitro assay, the ER-RUC provides direct interaction between chemical and ER. The assay provides consistent responses at the simple screening level, across laboratories, and these responses are in line with expectations for those chemicals tested whose ER binding behavior is well-established. Compounds that do not interact with ER consistently test negative in the assay. The assay is short and inexpensive compared to in vivo tests.

Limitations.

The assay is sensitive to many details of preparation and technique and can show wide variability if not performed exactly as stated in the protocol. It is subjected to problems if the receptor concentration in the cytosol is too low or too high. Another problem is that the tubes are not kept cold at all times during preparation, incubation and separation of bound from free tracer. This problem is certainly responsible of the the top plateaus for the standart chemicals (estradiol and norethynodrel) often exceeded the performance criteria by several tens of percentage points. However, the data

suggest that a lab that meets the performance criteria for teh standart and weak positive is likely to generate data that is much less varaiable than laboratories that do not meet performance criteria.

The analysis of datasets was relatively complicated. The standardization of the assay does not allow precise and replicable quantitative analysis of log(IC50) and RBA.

Finally, the assay requires the use of animals.

5. Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

Yes, the test substances were appropriately chosen to demonstrate the performance of the assay. The test substances correspond to a wide range of strenghs and chemical structure. They are 3 very strong, 3 strong (including E2), 3 moderate, 9 weak and 5 negative compounds. Among negative compounds, atrazine (non-binder but estrogen-active) was included.

Analytical methods and statistical methods were also appropriately chosen to demonstrate the performance of the assay.

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?

Considering the variability inherent in biological and chemical test methods, the results obtained with this assay are relatively repeatable and reproducible.

Very strong an strong compounds.

All of very strong, strong and moderate (17b-estradiol, 17-ethynylestradiol, DES, Meso-hexestrol, zearalenone, tamoxifen, norethynodrel, genistein and equol) binders were correctly determined to interact with rat uterine ER. The only exception is estradiol in laboratory Z. Binding at the concentrations tested (100 pM to 1 mM) showed clear interaction with the receptor but the laboratory did not adjust the test concentration range to a more dilute range that would have allowed characterization of the full binding curve as required by the protocol.

Note that laboratory Z misinterpretated (compare table 27 ER-Ruc and table 2 p22 appendix 8) strong affinity compound 1 (estradiol), 3 (DES) and 13 (tamoxifen). Note that tamoxifen is not by itself a strong binder. 4-hydroxy tamoxifen (a tamoxifen metabolite) is the strong binder.

When the three laboratories' RBAs were compared, similar values were obtained for very strong and strong compounds.

Moderate compounds

For moderate compounds, similar results were obtained for norethynodrel, equol and zearalenone while variability between laboratories is observed for genistein and tamoxifen (appendix 6 table 5, appendix 7 table 6 and appendix 8 table 2).

Weak binders

The weak chemicals had various classifications by the laboratories. Butyl paraben and bisphenol A were classified as positive by all of the laboratories.

Nonyphenol (mixture), 4-n-heptylphenol, and enterolactone were positive, negative and equivocal for laboratory X, Y, and Z, respectively. Kepone was positive in two laboratories and equivocal in the third. O,p'-DDT was negative in two laboratories and positive in the third. After analysis of the laboratory results, it was concluded in the ER-RUC ISR that these seven compounds were binders. I agree with the ER-RIC ISR exepted for o,p'DDT for which more experiments are necessary to conclude.

5alpha-dihydrotestosterone was negative in two laboratories and equivocal in the third.
Benz(a)anthracene was classified as negative by all of the laboratories.

Thus, most of the compounds that were expected to be weak binder were «positive« in their responses. However as mentionned above, the assay does not allow precise and replicable quantitative analysis of log(IC50) and RBA.

For the negative chemicals (corticosterone, progesterone, octyltriethoxysilane, atrazine and R1881),

laboratory X and Y had corticosterone listed as positive, and laboratory Z had it as equivocal. Curiously, this compound was correctly classified as negative in the optional portion.

Progesterone and octyltriethoxysilane were classified as negative by all three laboratories. Atrazine was classified correctly as negative by laboratory T and Z but positive by laboratory X.

R1881 that was designed to be a negative control for the assay proved to be positive for the control runs, and as a coded chemical it was positive for laboratory X and Y and equivocal for laboratory Z.

Thus, most of the compounds that were expected not to be binder were « negative» or « equivocal » in their responses.

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?

Yes, appropriate parameters were selected and reasonable values were chosen to ensure proper performance of the assay.

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

Yes, the data interpretation criteria are clear, comprehensive, and consistent with the stated purpose.

9. Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.

The estrogen receptor binding assay using rat uterine cytosol will represent a validated assay for simple screening for interaction with the estrogen receptor in the context of a battery of in vitro assays. EPA has optimized and standardized the most important parameters of this assay and has shown that the resulting protocol is transferable to others laboratories and enable to identify chemicals that have the potential to interact with the endocrine system.

.

Analysis of compounds binding curves in appendix II of Appendix 5-overall report, enable easily to identify the nature of the compounds.

Among very strong and strong binders (17b-estradiol, 17-ethynylestradiol, DES, meso-hexestrol, zearalenone and tamoxifen), zearalenone and tamoxifen are not strong binders (IC50s around 100 nM).

Among moderate binders (genistein, norethynodrel, equol), genistein (IC50 around 100 nM) is a stronger binder than norethynodrel an equol (IC50s around 1000 nM).

Among predicted weak binders (butyl paraben, nonylphenol, o,p'-DDT, 5alpha-dihdrosterone, bisphenol A, 4-n-heptylphenol, kepone, benz(a)thracene, enterolactone), it was very easy to identify real weak binders in the test (butyl paraben, nonylphenol, bisphenol A, 4-n-heptylphenol, kepone, enterolactone) and non binders (o,p'-DDT, 5alpha-dihydrosterone, benz(a)thracene).

Finally among predicted negative compounds (atrazine, corticosterone, octyltriethoxysilane, progesterone and R1881), it was easy to identify weak binders (corticosterone and R1881) and real negative compounds (atrazine, octyltriethoxysilane and progesterone). Curiously, corticosterone was negative when the tested optional chemicals. Concerning R1881, it is likely that the expectation of non-binding is not correct.

When the analysis was developed, the EPA expectation was that standardization of the assay would allow precise and replicable quantitative analysis of log(IC50)s and Relative Binding Affinities. The expectation was also that precise, standardized methods of analysis would contribute to reproducibility and therefore use in other applications such as structure-activity relationship models.

The variability of the results of the assay may not support for the moment use for quantitative structure-activity relationship model development. Even if strenght of binding could be evaluated, high variability of compounds RBAs were observed. In the face of the variability encountered, EPA is assessing whether such analysis could be replaced with a simpler analysis and still meet the needs of the Screening Program.

However, it should be remembered that while intralaboratory variability was disappointingly high for at least one laboratory in this study, such variability is not expected to be as much of a problem for laboratories that demonstrate the ability to meet the required performance criteria. The limited time available to run this large study on 23 chemicals apparently did not allow development of the proficiency necessary to obtain precise runs in all laboratories. The fact that results were almost all in accord with expectations when screening for interaction despite the variability in quantitative values shows that the assay is robust for this use.

Furthermore, use of recombinant receptor rather than receptor obtained from whole animals will enable to use purified receptor which would decrease the variability of the binding assay.

Finally, small ameliorations of the protocol, fine adjustement of performance criteria and increase of experience by the laboratories should allow the test to reach the expected purpose which is to determine the potential of a substance to interact with the endocrine system.

3.2 Ingemar Pongratz Review Comments

Review Questions and Answers
Ingemar Pongratz

1. Is the stated purpose of the assay clear?

Section 2, A.

In my opinion the test regarding the purpose of the assay should be modfied. The Estrogen Receptor Binding Assay is in itself only able to to identify compounds that bind to the estrogen receptors ERα and/or ERβ. However, compounds can interact with with the estrogen receptor system through alternative mechanisms that may not involve direct binding to the estrogen receptors.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

Section 2, B, C, D

Only partially. In my opinion the assay does not take into full account some issues. Recent experiments have shown that one of the estrogen receptor isoforms namely ERβ is under circadian control in the mouse. The circadian system is well conserved so it therefore likely that this is occurs also in the rat model.

Depending on the timepoint of cytosol preparation, the levels of ERβ expression may be very low.

These low ERβ levels may result in cytosol preparations that fail to detect compounds that preferentially interact with the ERβ isoform and thus may considered "safe".

This point should be taken into account to avoid missinterpretation of obtained results.

A second point regading the biolgical relevance of the assay is the role of the molecular chaperone hsp90 and its role in maintaining the estrogen receptors in a ligand binding state. Previous work with hsp90 associated receptors, in particular receptors like the GR, has demonstrated an important role for the hsp90 complex. In the case of the estrogen receptor ERα experiments have demonstrates that hsp90 is important but not crucial, for ligand binding. I am concerned that during cytosol preparation, the hsp90 complex may dissociate which would negatively impact on the receptors ligand binding activity. The presence of molybdate will stabilize the complex but it may not be sufficient.

In the case of the second estrogen receptor isoform ERβ, very little is know regarding the putative role of hsp90 and regarding the stability of the complex. Again this may cause problems and in particular may explain some of the interlaboratory variation.

3. **Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:**
    a) **comprehend the objective;**
    b) **conduct the assay;**
    c) **observe and measure prescribed endpoints;**
    d) **compile and prepare data for statistical analyses; and**

e) report the results?

What additional advice, if any, can be given regarding the protocol?

Section 2F

The protocol is technically sounds, and very detailed. In my opinion efforts should be done to make it easier to follow. It is stated that the groups performing the assays should have a suitable background in the field so they are well informed about the method. However I feel that the protocol is difficult to follow. In general I feel that an "Excecutive Summary" would be beneficial in particular to understand the purpose of the assay.

This comment is also relevant for point b; how to conduct the assay. Again a short annex would be improve clarity.

I have very limited experience regarding statistical analysis so I feel I am not competent to provide input regading this point.

4. Have the strengths and/or limitations of the assay been adequately addressed?

Section 2

I feel that the assay as first line approach is valid, but would benefit from additional scientific experimentation and validation, in particular characterize some of the issues regarding the circadian expression of the ERβ isofom and of the potential stability of the hsp90-ERα and hsp90 ERβ complex. I feel this limitations should be addressed in the text and efforts should be undertaken to fill the scientific gap regarding these issues.

A further limitation of the assay is due to the use of laboratory animals. There are a number of parameters that are difficult to control, such as quality of feed. Has the possibility to use established cell-lines been considered?

5. **Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?**

I think that both the test substances and the methods are appropiate for the purpose.
I have little experience with the statiscal methodology so have no opinion regarding the statistical methods that were chosen for the assay.

6. **Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?**

In general I would say yes. I would suggest that the experimental issues raised previously be taken into account.

The possible effect of the circadian system can be covered by characterizeng the perak of ERβ expression at different time points, the food composition of the animals should be centralized, and stabiltiy of the hsp90 complex needs to be better addressed. A experimental possibility would be to use cell-lines, I understand however that this would lead to less material being available.

7. **With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?**

In my opinion yes.

8. **Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?**

The data interpretation is in my opinion clear and consistent with the purpose.

9. **Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.**

I think the assay is suitable to identify compounds that bind to the estrogen receptors ERα and ERβ. Again, the assay should be regarded in the context of being one of several assays. I have however certain reservations to the statement that the assay will cover metabolism. Metabolism and the presence of the p450 enzymes is in many cases a cell and tissue specific process and all the necessary components may not be present in uterine tissue and cells.

3.3 Shlomo Sasson Review Comments

PEER REVIEW OF THE ESTROGEN RECEPTOR BINDING ASSAY AS A POTENTIAL SCREEN IN THE ENDOCRINE DISRUPTOR SCREENING PROGRAM

Shlomo Sasson
April 2009

1. **Is the stated purpose of the assay clear?**

   The rat uterine cytosol estrogen receptor (RUC-ER) competitive binding assay is one in a battery of assays aimed at providing validated strategy for the screening of endocrine disruptors (e.g., xenobiotics and environmental chemicals) that interact with the ER in target tissues and alter their normal functions. The purpose of the RUC-ER binding assay in this extensive project is clearly stated. The background information on the ER binding assay and the various options and choices of ER preparations from different animal species and recombinant human ER are well-presented. The strengths and shortcomings of the RUC-ER binding assay in view of numerous previous studies were taken into account in selecting and including this assay in the collection of other cell-based- and *in vivo* assays.

2. **Is the assay biologically and toxicologically relevant to the stated purpose?**

The stated purpose of the RUC-ER binding assay is clear- it aims at identifying compounds that interact with the ER by testing their potential to compete with the natural ligand, 17β-estradiol, for binding with rat ER. It is clearly acknowledged that this competitive binding assay is not aimed at ascertaining the functional properties (weak/strong agonists, partial/mixed agonists or antagonists) of the test chemicals. However, the goal to categorize these molecules according to their binding affinities with the receptor is feasible and suits the formal objectives of the assay. It should be noted, however, that while there is a high degree of confidence that high affinity binders/competitors may activate or inhibit ER function in *in vitro* and *in vivo* models, it is doubtful whether weak or very weak binders/competitors would substantially interact with the ER and exert biological functions *in vivo*. Thus, this RUC-ER assay (and most likely the hER assay that is evaluated in parallel) allows the classification of potential endocrine disruptors by virtue of their intrinsic binding affinity for the ER. Further analysis of the biological or toxicological effects of these compounds entails independent cell- and animal-based assays.

3. **Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:**

a) **comprehend the objective**;

The objective of the RUC-ER binding assay is clear. The background information on the assay and its inclusion in the study are well-discussed. The assay was developed following an in-depth review of the literature, which provides the scientific basis for the assay. The standardization and attempts to optimize the assay following the First Interlaboratory Study resulted in detailed, comprehensive and clear protocols. Experienced laboratories, such as the 3 that participated in the Second Interlaboratory Study, are expected to adapt and conduct the assay with minimal deviations from the protocols. Nevertheless, the results obtained from these 3 laboratories and their comments on problems encountered in the course of the study do indicate the need for further examination and optimization of the protocols.

b) **conduct the assay;**

The conduct of the assay is expected to follow the detailed protocols. The two-test approach (the [^{3}H]-17β-estradiol saturation binding analysis, on one hand, and the competitive binding assay, on the other) is very important as it provides intra- and inter-laboratory quality controls and intrinsic assay controls. There are, however, certain problems that were encountered 'on the bench' or when data were collected, analyzed and interpreted. These issues are presented and discussed at the end of this section under **f**.

c) **observe and measure prescribed endpoints;**

The protocols are clear, detailed and meticulously describe all necessary technical aspects- from uteri excision to data acquisition and analyses. By conducting the assay according to protocols each laboratory is expected to observe and measure prescribed endpoints. In some cases along the Second Interlaboratory Validation Assay not all endpoints were achieved- mostly due to deviations from the prescribed protocols and/or due to technical problems.

d) **compile and prepare data for statistical analyses;**

The platform provided to compile and prepare the data for kinetic analyses followed by statistical analyses has served well the participating laboratories and suits the requirements of the assay. Not being an expert in statistics, I cannot comment on the statistical methodology (Appendices 4 and 9). However, Appendix 13, which depicts graphs of acceptable saturation and competitive binding curves, raises some reservations on the acceptability of these graphs: the range of 17β-estradiol IC_{50} values (TRL's and Hammer's curves, Appendix 13/pp. 2-3) and of the weak positive runs (*ibid*, pp. 5-6) covers nearly an entire order of magnitude. It appears that narrower pre-defined limits are required in these statistical analyses.

e) **report the results?**

All reports of the results by the three laboratories are extensive and provide a solid basis for intra- and interlaboratory analyses of the data. Attention should be given to the

suggestions related to **improving initial data manipulations** and **curve fitting** (Appendix 5/pp. 16-17).

f) **What additional advice, if any, can be given regarding the protocol?**

1. The choice of ovariectomized Sprague-Dawley rats for uteri collection is good. However, the argument made by the RTI scientists to extend the 8-day period after ovariectomy before excision of uteri due to remaining endogenous estrogens (Appendix 5/p.5) is not pharmacokinitically sound: the half-life of 17β-estradiol in Sprague-Dawley rats is about 10 hours (*Petroff and Mizing. Reproductive Biology, 3:131, 2003*). Thus, a full clearance of 17β-estradiol is expected after nearly 20 half-lives within the 8 days following ovariectomy. It is more likely, however, that the presence of stimulated uteri (due to fluid retention) was the consequence of an incomplete ovariectomy. In fact, RTI scientists confirmed the presence of residual ovarian tissue in some operated rats (*ibid*). It is therefore suggested that in cases where such remnants of ovarian tissue are found or when the excised uteri appear imbibed to discard them from the assay.

2. The requirement to perform a saturation binding assay for each RUC preparation prior to the competitive binding assay is mandatory: it allows the determination of the ER concentration in cytosols in term of maximal binding capacity of [^{3}H]-17β-estradiol. The large variation in the total protein content of the various rat uterine cytosols prepared by the 3 participating laboratories, particularly in the three cytosols prepared in Lab Y (ISR, Table 22, p.53) attests to this need. Despite the extraordinary high protein content in 2 of the cytosols in Lab Y, the mean maximal binding (B*max*) reported was 2-3-fold lower that that reported by Lab X, whose cytosols contained significantly lower protein content. This is probably due to the addition of a fixed amount of cytosolic proteins to the test tubes, leading inevitably to low content of ER. Therefore, I suggest to pre-determine the range of acceptable protein content in rat uterine cytosols. Large variations in protein content as reported by participating laboratories may complicate the interpretation of results.

3. It is important to keep the total receptor concentration in the binding assay low enough to restrict the binding interaction of 17β-estradiol with the receptor to a simple (single and non-interacting binding sites) mechanism (linear Scatchard plot, Hill coefficient=1) and to avoid complex binding kinetics due to receptor dimerization that occur at higher receptor concentration. It is therefore recommended to scrutinize all experiments where the Hill coefficient of 17β-estradiol binding is significantly higher than 1.0 and dilute the cytosol accordingly. Dilution of the ER in the competitive binding assay to 0.5 nM usually suffices the requirement of a simple binding kinetics.

4. Often, the 'top plateau levels' of competition curves were markedly higher than the expected 100%. There are several technical explanations to these results, such as, underestimation of the control binding, receptor instability or variable concentration of solvents in test tubes. The recommendation to use a fixed volume of solvent (2 % ethanol) in all tubes, including controls, will eliminate the latter cause. It is suggested to design and include 'receptor stability' controls in all assays. The simplest assay requires an incubation of the cytosol under the same assay conditions of the competitive binding assay (16-20 hours at 4ºC) with no ligands. This cytosol is then used in a binding assay along with a freshly thawed cytosol. The binding capacity of these two cytosols is compared in a binding assay with saturating concentrations of [^{3}H]-17β-estradiol, an efficient anti-protease cocktail and incubation at high temperature (15-20ºC) for 1-2 hours followed by the HAP treatment. This simple assay eliminates misinterpretation of data due to receptor instability during the long incubation period at 4ºC.

5. By definition, if the [^{3}H]-17β-estradiol is present at a concentration that saturates the receptor- an equal concentration of unlabeled 17β-estradiol is expected to reduce the specific binding of [^{3}H]-17β-estradiol by 50% due to the 1:1 dilution of the radioactive ligand. In the case of the presence of

excessive binding sites (non-saturating conditions) the 50% competition of the labeled 17β-estradiol with the unlabelled 17β-estratiol is observed at a ratio higher than 1:1 for the unlabeled estradiol. This may lead to an inaccurate estimation of IC_{50} values. The range of IC_{50} values is depicted in the Acceptable Standard 17β-estradiol Curves (Appendix 13/pp.1-3): the range reported by TRI is narrow enough and agrees with the assay's terms. However, the ranges shown by TRL and Hammer cover an entire order of magnitude. These wide ranges most probably result from variable receptor levels and non-saturating [^{3}H]-17β-estradiol concentrations. Therefore, I suggest using the data obtained from the standard 17β-estradiol saturation curves to calculate and use the optimal receptor concentration for maximal binding capacity (saturation) of [^{3}H]-17β-estradiol. In Appendix 6 (Appendix D, p.3, 8 and 12) Lab X presents ER binding curves that were saturates with 2-3 nM 17β-estradiol. Lab Y presents a summary of 3 saturation curves in Appendix 7 (p.16) in which maximal binding was observed with 0.5-1.0 nM [^{3}H]-17β-estradiol. Lab Z gives 4 saturation curves (Appendix 8/p. A-17, , C-20, D-17, E-17) which did not reach saturation at 1-2 nM [^{3}H]-17β-estradiol and one curve (B-22) that saturated between 1.3-3.6 nM [^{3}H]-17β-estradiol. These inconsistent data may explain intra- and interlaboratory variations of the calculated kinetic parameters (K*d*, IC_{50}, and RBA). The equivocal results of the competitive binding assay of Test Chemical #1 (17β-estradiol) presented by Lab Z could have been avoided if the assays were conducted with a prescribed concentration ER and saturating concentrations [^{3}H]-17β-estradiol.

6. It is highly recommended that researchers will assess potential precipitation of hydrophobic test compounds in the binding assay mixture by using physical methods such as light scattering spectroscopy. Precipitation problems can often be solved with some organic solvents. Appendix 3 shows the compatibility of the binding assay to 2% ethanol. However, the recommendation to allow DMSO up to 20% in the assay is worrying; the U-shaped displacement curves of various ligands to the ER (Figures16, 18, 22

and 23) indicate that DMSO may affect the receptor and change its binding kinetic properties in a concentration-dependent manner. Indeed, such effects of solvents (*i.e*., dimethylformamide) on the binding kinetics of the ER were previously reported (*Sasson & Notides, J. Steroid Biochem. 29:491-5, 1988*).

7. The absorption of hydrophobic chemicals to test tube walls is not always eliminated by using siliconized glass (borosilicate) tubes. Compounds with a high partition coefficient (*e.g.,* the Log*P* of tamoxifen is 6.58 in comparison with 3.67 of 17β-estradiol) do interact with these seemingly 'inert' surfaces. This absorption problem may be solved by other means than just using solvents at different concentration. It was shown (*ibid*) that a protein content higher than 2.5 mg/ml in the binding assay reduced significantly tamoxifen lose to the tube walls. Various proteins, such as insulin or IgG, which minimally increase the non-specific binding of 17β-estradiol, can be added to low-protein cytosols for this purpose. Yet, such unusual solutions are rare because ethanol is a good solvent for most chemicals and is well tolerated in the binding assay up to a final volume of 2%.

8. The HAP sedimentation procedure is used for separating free from bound [^{3}H]-17β-estradiol. Following washes the HAP slurry is extracted with ethanol and a 1 ml aliquot is then mixed with 14 ml of scintillation cocktail to measure radioactivity in a β-counter. An accurate conversion of counts-per-minute (cpm) to disintegrations-per-minute (dpm) is of a paramount importance for subsequent data analysis. The counting efficiency may differ among various scintillation cocktails and β-counter settings. Various methods are available to estimate counting efficiency and correctly calculate dpm: internal standards, Channel Ratio (CR) or external standard (γ-radiation). It is imperative that each participating laboratory be able to reliably calculate dpm of radioactive samples. The use of internal standards (that is, fixed amount of μCi of [^{3}H]-17β-estradiol counted with 1 ml ethanol in 14 ml of scintillation cocktail) is often adequate; Erratic CR values identify counting vials in which quenching was irregular.

9. RTI scientists listed several suggestions regarding the assay protocol (Appendix 5/p. 15). The suggestion to **conduct an abbreviated pilot study** (screen test compounds at 10 μM and further characterize only those that inhibit [^{3}H]-17β-estradiol binding by 50% or more) sounds reasonable. However, the limit in my opinion should be lower than 50% to eliminate false negative determinations from the assay. The suggestion to **determine the K*d* values of [^{3}H]-17β-estradiol only once** for each newly prepared cytosol is acceptable, as long as the B*max* values do not significantly change when the cytosol is tested repeatedly in independent assays. Moreover, inclusion of a control for receptor stability (see above) will also warrant similar experimental conditions among assays. The suggestion **not to run full standard assays** (17β-estradiol and norethynodrel), but use the compounds at their respective IC_{20} and IC_{80} values is problematic: close examinations of the acceptable standard curves in various runs of each participating laboratory and among them fails to provide absolute concentrations of each compound that induce 20 or 80% displacement of [^{3}H]-17β-estradiol binding. Furthermore, full standard curves and the experimental values of IC_{50} of unlabelled 17β-estradiol do provide important information on the quality and saturability of the assay. Similarly, a Hill coefficient value close to 1 indicates a simple binding mechanism. I do, however, tend to agree that the norethynodrel standard curve is redundant. **Stability of materials** is very important and I fully support the suggestions regarding this issue. Similarly, problems with **solubility of chemicals** are critical and the suggestion to employ various approaches to solubilize the compounds, other than these specified in the protocols, is sound. Yet, this requires control experiments showing that the solubilizing solvent or other compounds used do not interfere with the binding kinetics of 17β-estradiol. Obviously, **brands of reagents and suppliers** do vary. Thus, the idea to provide a "suggested supply list" is encouraged. The problem regarding **difficulties in pipetting 10 μL of test chemical and standards** can easily be solved if these test chemicals or standards are pipetted first to dry tube- no immersion of pipette tips in the binding mixture

or an incomplete delivery of the tip's content are encountered this way. Accordingly, the order of addition of the assay components should be modified accordingly.

10. The case of dual (mixed) estrogenic-antiestrogenic properties of certain ER ligands was not addressed in the study. Some compounds, like estriol, act as weak estrogens when present or administered alone to immature or ovariectomized animals. However, when present in excess over estradiol, these compounds exert potent antiestrogenic properties. This duality in function depends on the concentration of the competing ligand relative to the ambient 17β-estradiol concentration. This phenomenon results from a partial activation of the ER when the test compound is present alone, and to an aberrant dimrization of the ER in the presence of the natural ligand (*Melamed et al., Mol. Endocrinol. 11:1868-78, 1997*). The simple binding assay employed in this study is not aimed at analyzing such complex kinetic interactions. However, these complex interactions may occur in animal studies. Hence, it is recommended that functional assays aimed at determining the properties of ER binders *in vitro* and *in vivo* would carefully explore their potential to act as mixed agonists/antagonists. Equally important is the potential of such compounds to exert antiestrogenic properties in female with normal ovary function and secretion of 17β-estradiol, while acting as estrogenic agonists in males, due to the absence of significant levels of natural estrogens. This spectrum of functions should be acknowledged and examined carefully in studies that aim at determining biological and toxicological functions of such ER ligands.

4. Have the strengths and/or limitations of the assay been adequately addressed?

The strengths of the assay are adequately addressed (ISR, pp.68-9). The major strength was the ability of the RUC-ER assay to classify correctly most test compounds, despite inconsistencies among the participating laboratories and various technical problems. Clearly, once the recombinant hER binding assay is completed, a thorough comparison and analysis of the results of the two independent assays is required for further validation

of the RUC-ER as a relatively simple and affordable screening assay. The weaknesses of the assay (*ibid*, p.69-70) refer mostly to technical and methodological aspects such as, ER concentration in the assay, insolubility of test compounds, complicated analyses of data and the inevitable use of rats. There are no satisfactory explanations to the lack of adherence of some participating laboratories to the standard protocols.

5. **Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?**
   Test substances were carefully collected to cover a wide range of compounds that bind to the ER at very strong (higher than 17β-estradiol), moderate or weak affinities, regardless of their biological functions. The analytical methods are appropriate to demonstrate the performance of the assay. As stated above, I refrain from commenting on the statistical methods.

6. **Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?**
   In most cases the answer to this question is positive. The summary Table (Table 27, ISR, p.60) shows that despite a considerable variability in the data obtained by the 3 laboratories, most test compounds were correctly classified. The case of Compound #1 (17β-estradiol) in Lab Z that was classified 'equivocal' is peculiar and most probably results from deviations from the standard protocol of the competitive binding assay. In fact, it is somewhat surprising that despite considerable variability in datasets obtain by participating laboratories and among them- the final evaluation was powerful enough to categorize most test chemicals according to their rank. Yet, adherence to the protocols and further optimization of the assay will eventually lead to more reproducible results.

7. **With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?**
   The saturation binding assay of 17β-estradiol and the competitive binding assay were introduced some 40 years ago with the discovery of the estrogen receptor and the availability of radioactive estradiol. The kinetic parameters, K*d*, B*max* and RBA have been standard endpoints in numerous binding assays. The K*d* values for 17β-estradiol

binding to the RUC-ER in this study fall within the acceptable range. Obviously, B*max* values can vary widely, depending on the quality and ER concentration of the uterine cytosols.

8. **Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?**

   The data interpretation follows rigorous kinetic and statistical analyses. Both provide clear representation and analyses of the data and are fully consistent with the stated purpose of the assay.

9. **Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.**

   This assay is one in a battery of assays aimed at identifying endocrine disruptors. The RUC-ER binding assay, which was developed some 40 years ago, provides simple kinetic analyses of ER binding interactions with natural and synthetic agonists and antagonists. The principles of the assay have not significantly changed over this period. A major development is the source of the ER: uteri from various species (mostly rats and calves) have been used extensively for years. The current availability of recombinant human ERα and ERβ now enables studies more relevant to human. Thus, it will be interesting to compare the results of the RUC-ER assay to the recombinant hER project that is run in parallel. Both assays are relatively simple and can be included in the battery of assays to provide a first-line screening of potential ligands to the ER. I believe that the RUC-ER meets these criteria. Nevertheless, the protocols of the assay require some further optimization.

3.4 Marie-Louise Scippo Review Comments

PEER REVIEW OF THE ESTROGEN RECEPTOR (ER) BINDING ASSAY
April 2009

1. Is the stated purpose of the assay clear?

Yes. It is clearly stated that the estrogen receptor rat uterus cytosol (ER-RUC) binding assay is one of the tests of a battery of complementary screens, included in the endocrine disruptor screening program Tier-1 battery.

It is also clearly stated that the aim of the ER-RUC is to detect an interaction with the estrogen receptor, not to identify the mechanism of action (stated i.e. p.8, 65, 69-70 of the ISR, and page 5 of Appendix 1).

What is less clear is the weight to give to the result obtained for an unknown chemical using the ER-RUC assay (interactive or not with the ER) within the battery of the Tier-1 program.

It should be interesting to give, in the introduction of the integrated summary report (ISR) (page 2, under C. "*The Tier1 battery of assays*"), a description of the strategy that will be used to classify a chemical as negative or positive after the Tier 1 screening, that includes various *in vitro* and *in vivo* assays (ISR, page 3, table 1), and to give the weight of each assay in the final decision of the Tier 1 screening.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

Yes, it is, even if this single assay gives no indication about the toxicity of a chemical.

The stated purpose of assays involving estrogen receptors is to evidence an estrogenic or an anti-estrogenic activity of the test chemicals.

For both estrogenic and anti-estrogenic compounds, in most cases, the first step of their biological activity is the binding to the estrogen receptor (ER), before target genes transcriptional activation (estrogenic compounds) or inhibition of it (anti-estrogenic

compounds). There are two isoforms α and β of the estrogen receptor, and some compounds act more specifically on one or another form.

As the rat uterus tissue displays both α and β isoforms, the ER-RUC allows detecting both α and β ER ligands, which is an advantage over the binding assays using a recombinant receptor of a single isoform.

As the transcriptional activation assay of the Tier-1 battery is specific to the α isoform of the receptor, it is relevant to use the ER-RUC allowing the detection of the binding to both α and β isoforms, in a screening approach using a battery of complementary assays, in order to decrease the rate of false positive results of the Tier-1 testing.

3. **Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:**
   a. **comprehend the objective;**
   b. **conduct the assay;**
   c. **observe and measure prescribed endpoints;**
   d. **compile and prepare data for statistical analyses; and**
   e. **report the results?**

What additional advice, if any, can be given regarding the protocol?

The answer to questions a) to e) is globally YES, but some clarifications are needed (see here below).

1. Data for plotting the competitive binding curves : the y axis.

In Appendix 1, page 46, paragraph 10.7, the data analysis is not totally clearly explained.

It is indicated that "*The competitive binding curve is plotted as specific [^{3}H]-17β-estradiol binding versus the concentration (log10 units) of the competitor*."

If we look at 10.7.1. "*Terminology*", the definition of the specific binding is "*Total binding minus non-specific binding*", but the "*total binding*" is not defined, and thus the way to calculate the specific binding is not indicated.

It should be added that :
Total binding is radioactivity in DPMs in the tube that contains [3H]-17β-estradiol and receptor, in the absence or presence of competitor.

Furthermore, it is not clear which figures we have to plot on the y axis of the competitive binding curve.

In Appendix 4, page 14, Paragraph 2.1.1. (Input data specification), it is indicated that "*input data for the dependent variables should be standardized and expressed as "% binding of the reference ligand to the receptor*", which I translate as the ratio :

$$\frac{\text{Total binding (in presence of competitor) - NSB}}{\text{Total binding (in absence of competitor) - NSB}}$$

As this is indicated nowhere in the documents (may be it is clear on the excel worksheets, but unfortunately, we didn't receive the excel files, this should be added to the terminology, under "*specific binding*".

2. Receptor concentration in the assay
2.1. Performance criteria of the saturation binding assay: Standardization of the receptor concentration.

In Appendix 1, page 20, 9.1.5. "*Standardization of receptor concentration*", it is indicated that "*For the saturation assay, the optimal protein concentration binds 25 - 35% of the total radiolabeled estradiol that has been added to the tube. To ensure that this percent range of radioligand is bound at the lowest concentration of radioligand added to the assay, the 0.03 nM concentration shall be used to make this determination for the saturation binding assay*".

Under 9.7. ("*test report*", pages 26 - 27 of Appendix 1), it is not clearly asked to the lab to report the percent of bound radioligand at the 0.03 M concentration.
It should be clearly asked to the labs to report this value.

Furthermore, I tried to retrieve these data in the raw data of the individual labs, and I found for both lab Z and Y (Appendix 8, page A-2 and Appendix 7, page B-9 respectively), that the percent of radioligand bound at 0.03 M radioinert estradiol was indicated in a column named "*10 Percent rule*", and the criteria was obviously that the percent of radioligand bound should be below 10%. This is not in agreement with the protocol (which requires for 25 - 35 %).

2.2. Performance criteria of the saturation binding assay: Standardization of the receptor concentration.

In Appendix 1, Page 47, 10.7.3. "*Performance criteria for the competitive binding assay*", it is indicated that: "*Ligand depletion is minimal. Specifically, the ratio of total binding in the absence of competitor to the total amount of [^{3}H]-17β-estradiol added per assay tube is no greater than approximately 15%*".

Again here, it is not asked to the labs to report this ratio, but only the amount of protein added per tube.

It should be asked to the lab to report this parameter too.

Again here, lab Y used the below "*10% rule*", and generally, the percent of radioligand bound was around 3%, which I find too low (Appendix 7, page D-3).

I was not able to find the information in the report of lab X.

Furthermore, I see here a contradiction with Appendix 1, page 20, 9.1.5. where it is said that "*For the saturation assay, the optimal protein concentration binds 25 -35% of the total radiolabeled estradiol that has been added to the tube*". This would mean

that a different protein concentration should be used for saturation binding assays (25 - 35% binding of the total radioligand added) and for competitive binding assays (15% binding of the total radioligand added), which is not what is done in practice. Please, clarify that point in the protocol.

3. In a validation study, the protocol given to the participating laboratories should be a definitive one. As several changes have been made in the protocol after the second interlaboratory study (we see that by comparing the modified protocol described in Appendix 1 and the original protocol described in Appendix 5), it should be relevant to organize a third interlaboratory study using (and reviewed) a definitive protocol.

A specific remark :
Appendix 1, page 48, table 9 : the lower limit of the top plateau for estradiol is 90% and not 94%.

4. Have the strengths and/or limitations of the assay been adequately addressed?

The strengths and the weaknesses have been adequately addressed.

The strengths of the assay are, for the most, relevant for all estrogen receptor binding assays but one important specificity of the ER-RUC is stressed : it is that the rat uterine cytosol contains both isoforms alpha and beta of the estrogen receptor. The high degree of homology between the rat and the human ER ligand binding domain could also be mentioned here.

The main weakness of the assay comes from the use of animals to prepare the binding fraction, with not only ethical, but also technical consequences (lack of reproducibility in receptor preparations). The solution is the use of recombinant ER binding assays, using both alpha and beta isoforms of the ER.

If a third interlaboratory study is organized, it would be interesting to analyze the same list of chemicals with both ER-RUC and recombinant ER binding assays. The comparison could provide different conclusions that could add some weight to the strength of the ER-RUC

using both α and β isoforms of ER compared to a recombinant ER binding assay using the α isoform only.

5. **Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?**

The choice of a panel of strong, very strong, moderate, weak or negative binders seems relevant.

However, it would have been better to avoid substances without available data in the literature about their binding capacity (in ER-RUC assays) to the ER.

In ISR, page 52, table 21, page 52, references of "*historical*" RBAs should be given.

6. **Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?**

The results were sufficiently repeatable and reproducible for the reference compounds (estradiol and norethynodrel), but for test chemicals, some results were too much dispersed (for example, for test chemical n°1, for lab Z, see in Appendix 5, page 1 of Appendix 2).

7. **With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?**

Performance criteria were relevant, but I would add some criteria such as those presented in the answer to question n°3.

To ensure the performance of the assay to detect ER interactive chemicals, I would add some criteria such as maximum rate of false negative and false positive decisions, measured with known substances.

In this validation study, 23 test chemicals were used to check this parameter, but the results are not very conclusive. On 22 chemicals (if we exclude R1881), respectively 5, 2 and 5 false decisions were taken by labs X, Y and Z (see ISR, page 60, table 27).

This rate of false decision will also depend of the weight of the ER-RUC assay in the Tier 1 screening (see my answer to question 1).

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

Yes, but there are some inconsistencies between the different documents.

Appendix 5 presents the report of the second inter-laboratory validation of the estrogen receptor binding assay (rat uterin cytosol). First, in Appendix 5, labs are named A, B, C, which is confusing for the reader, because in the ISR, the labs are named X, Z and Y respectively.

In Appendix 5, page 10, it is indicated that the labs were required to classify the unknown chemicals as positive, negative or equivocal binders. This is not the classification described in the protocol (Appendix 1, page 48,10.7.4), where the classification is described as interactive, not interactive or equivocal.

The results of the classification of the 23 test chemicals by the 3 labs X, Y and Z are presented in both ISR and Appendix 5 documents, in table 27 and 5 respectively. I expected that the results were the same in both documents, but they are not for 14 chemicals out of 23, without any explanation in the ISR. This is very confusing for the reader!

9. Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.

The introduction clearly indicates that the ER-RUC assay is one of the complementary assays included in Tier-1, listed in table 1, page 3. In page 1, it is indicated that "*A negative result in Tier 1 would be sufficient to put a chemical aside as having low to no potential to cause endocrine disruption, whereas a positive result would require further testing in Tier 2* ". And in page 7, we read : "*An individual assay may serve to strengthen the weight of evidence in a determination (e.g., positive results in an ER binding assay in conjunction with positive results in the uterotropic and pubertal female assays would provide a consistent signal for*

estrogenicity)". However, it is not clear, how is taken the final decision after Tier-1 (negative result or positive result), if all the complementary assays of the Tier-1 battery don't give concordant results (for example positive results in the ER binding assay but negative result in the in vitro transcriptional assay).

In page 2 of the ISR, there is a mention about "*false negatives*" and "*false positives*" resulting from the Tier-1 screening in the sentence "*Maximum sensitivity to minimize false negatives while permitting an as yet undetermined, but acceptable, level of false positives*". It would be necessary to explain how a positive or a negative decision is taken after the Tier-1 screening, or to mention a reference giving this explanation.

Additional remarks about the integrated summary report (file ER-RUC ISR, v3.16c.doc)

- Please, add a list of acronyms and abbreviations, it would be easier for the reader.
- Page 7, the reference "*USEPA, 2008*" is indicated in blue and underlined, but there is no hyperlink.
- Page 9, last paragraph of "*V. Overview of the assay*": In the sentence: "*The competitive assay measures the binding of [^{3}H]-estradiol at a fixed concentration in the presence of a wide range (eight orders of magnitude) of test chemical concentration*". Eight orders of magnitude should be replaced by seven orders of magnitude since the concentrations tested are from 100 pM to 1 mM (see section F, page 16, and Appendix 1).
- Page 10: in the sentence: "*The solvent used for a test chemical must also be used for the reference chemical (inert 17β-estradiol) and the control chemicals (norethynodrel and octyltriethoxysilane) unless the solvent is water*." replace "*inert 17β-estradiol*" by "radioinert 17β-estradiol".
- Same remark for the table in page 12.
- Page 15, in the sentence: "*If the test substance interacts with the receptor, it inhibits the binding of increasing amounts of radiolabeled estradiol*." please delete "*increasing amounts*", because it is a fixed amount of radiolabeled estradiol.
- Page 19, table 3: "*Performance criteria for competitive binding, reference and weak positive controls*". The figures of table 3 are different from figures of table 16 :

"*Performance criteria for second interlaboratory study*". Is table 3 valid for the 1st interlaboratory study only? As I understood, octyltriethoxysilane was not used as a negative control during the 1st interlaboratory. And in table 16, R1881 is still mentioned as negative control. Please give explanations about the differences between table 3 and table 16. Furthermore, the upper limit for the top plateau level for norethynodrel (% binding) is 110 and not 10.

- Page 20, paragraph III.A. "*Buffer Composition and receptor concentration*". This paragraph presents the evaluation of the effect of the receptor concentration, during the optimization of the assay. It is concluded that the optimal concentration is 50 µg protein/tube. Nothing is clearly said about the importance of the percentage of binding of the radioligand. It would be clearer to indicate, as it is in the protocol (Appendix 1, page 20, paragraph 9.1.5.), that the optimal receptor concentration has to be adjusted to a quantity of protein that "*binds 25 -35% of the total radiolabeled estradiol that has been added to the tube*" (measured using 0.3 nM estradiol in saturation binding assays). Furthermore, the protocol (Appendix 1, page 25) indicates that the protein concentration corresponding to that binding is generally in the range between 35 and 100µg/tube. For competitive binding assays, the protocol (Appendix 1, page 33) indicates that "*the optimal amount of cytosolic protein added should contain enough receptor to bind no more than 10 - 15% of the radiolabeled estradiol that has been added to the tube*".
- Page 20: In the sentence (last sentence of the page): "*The original design called for 100 mg of protein per assay tube. Protein concentrations from 25 to 100 mg/tube were evaluated*.", please consider to replace "*mg*" per µg.
- Page 21: In the sentence : "*The dpm values of the 100% binding tubes in these assays ranged from 1 to 4 percent of those of the hot tubes, indicating that ligand depletion was not significant*.", please define what are the "*hot tubes*" and preferably don't use "*hot tubes*", which is laboratory "jargon".
- Pages 27 and 28, Figures 3 and 4 are not cited in the text.
- Page 45, under "*B. Second interlaboratory study*", in the sentence: "*Therefore in the second study one stock solution was made in solvent and this solution diluted sequentially with buffer*.", please, consider to replace "*buffer*" by "solvent".

- Page 46, table 16 is still mentioning R1881, even if it was finally not used as a negative control.
- Page 47, table 18, Lab X reported values for the "*Ki*", but nothing is said in the text about the Ki. It would be better to delete this row from the table, and if not, to give some comments about the Ki.
- Pages 48 and 49, table 19 and 20 show results of qualification runs of the 3 labs participating to the 2nd interlaboratory study. The results are compared with performance criteria described in table 16, and non compliant results are highlighted. In table 16, for both estradiol and norethynodrel, lower and and upper limits for bottom plateau level are -5.0 and 1.0 % binding, respectively and lower and and upper limits for top plateau level are 90.0 and 110.0 % binding, respectively. In tables 19 and 20, cells containing figures not corresponding to these criteria are highlighted, but there are some inconsistencies (some cells should be highlighted and are not and others are highlighted and should not be), please check this.

3.5 William Welsh Review Comments

Charge Questions: Response to each of the following questions:

1. Is the stated purpose of the assay clear?

   The assay is described in a succinct and clear manner. The word "specificity", which is used on page 8 with respect to the saturation binding assay, should be defined in this context.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

   The assay is relevant to the stated purpose, i.e., to determine the ability of a compound to interact with the ERs isolated from rat uteri". As discussed in the ISR on page 64, it is not clear whether testing compounds at 1 milliMolar (1 mM) is toxicologically relevant. Although the rationale for testing compounds at this high, physiologically non-relevant concentration is explained adequately (pp. 64-65), the results obtained may be difficult to explain and subject to mis-interpretation and even false conclusions.

3. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:
   - **a)** comprehend the objective;
   - **b)** conduct the assay;
   - **c)** observe and measure prescribed endpoints;

d) compile and prepare data for statistical analyses; and
e) report the results?

What additional advice, if any, can be given regarding the protocol?

The protocol adequately describes the assay in a clear and concise manner. In terms of advice, a *glossary of terms* would facilitate clarity between the EPA and the laboratory and to help avoid any misunderstanding or mis-interpretation in meaning of specific terms described in the protocol(s).

4. Have the strengths and/or limitations of the assay been adequately addressed?

   The strengths and limitations of the assay are not clearly specified the ISR. This should be done.

5. Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

   Yes, these three aspects have been amply presented in the ISR. The substances and methods are appropriate to demonstrate the performance of the assay.

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?

   Yes, the results from the assay were extremely repeatable and reproducible especially in view of the variability inherent in biological and chemical test methods.

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?

   Yes, appropriate performance criteria were selected to ensure proper performance of the assay. These criteria are clearly articulated throughout the ISR.

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

   Yes, the criteria for data interpretation are described in a clear, comprehensive and consistent manner.

9. Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.

   The ER binding assay is fairly adequate as an initial screening tool. However, the ISR should describe to what extent the results (and/or interpretation of the results) of the assay might be influenced by recognition of non-genomic estrogenic signaling effects.

Appendix A

CHARGE TO PEER REVIEWERS

CHARGE TO PEER REVIEWERS
for
INDEPENDENT PEER REVIEW OF THE ESTROGEN RECEPTOR BINDING ASSAY AS A POTENTIAL SCREEN IN THE ENDOCRINE DISRUPTOR SCREENING PROGRAM (EDSP) TIER-1 BATTERY
March 13, 2009

Background:
Section 408(p) of the Federal Food Drug and Cosmetic Act requires the EPA to:

develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or other such endocrine effect as the Administrator may designate [21 U.S.C. 346a(p)].

EPA established the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), a panel of scientists and stakeholders, to provide recommendations on how to implement the Endocrine Disruptor Screening Program (EDSP). EDSTAC recommended expanding the EDSP beyond the estrogens to include the androgen and thyroid hormone systems, and beyond humans to include wildlife. EDSTAC also recommended that several assays be included in a Tier 1 battery of assays to determine the potential of a substance to interact with the endocrine system, and several additional assays to be included in a Tier 2 battery of assays to confirm interaction, establish whether adverse effects may occur, and examine the dose-response relationship.

One of the test systems recommended by the EDSTAC for the Tier 1 battery was the estrogen receptor (ER) binding assay. Its purpose in the Tier-1 battery is to detect chemicals that may affect the endocrine system by binding to the ER. EPA requested the National Institute of Environmental Health Sciences (NIEHS) to prepare a comprehensive historical review and critical evaluation of ER binding assays that had been reported in the scientific literature. According to the Expert Panel convened by NIEHS which examined the review, no existing ER binding method was adequately detailed and standardized to be considered "validated". The Expert Panel recommended that the EPA focus its attention on validating an ER binding assay that uses recombinant receptor rather than receptor obtained from whole animals (rat uterine cytosol, RUC). The EPA subsequently initiated such a validation effort for a human recombinant ER (hrER) binding assay, but felt that it had already invested so much time and effort in validating the ER-RUC assay that it would be appropriate to complete the validation of that assay as well. The current peer review focuses on the ER-RUC assay; peer review of the hrER assay will be conducted separately since validation of that assay is running somewhat behind that of the ER-RUC assay.

Although the ER binding assay, like all of the assays in the Endocrine Disruptor Screening Program, will be peer reviewed separately from the other assays in the program, it is expected to be used in conjunction with other assays to determine the potential of a chemical to interact with the endocrine system. This "battery of assays" approach was designed by EDSTAC to take advantage of the unmediated sensitivity of *in vitro* assays and other pathway-specific assays, while at the same time incorporating whole-animal assays that can detect effects that may be due to metabolites and some of which can detect effects from several different endocrine-related

pathways. A weight-of– evidence approach using information from all of the assays, not just the ER binding assay, will be used to determine whether a chemical substance has the potential to interact with the endocrine system.

This peer review should focus on the strengths and weaknesses of the ER binding assay itself, not on the Integrated Summary Report *per se* or any of the individual studies that were conducted as part of the validation of the ER binding assay.

Charge Questions:

Please respond to each of the following questions:

1. Is the stated purpose of the assay clear?

2. Is the assay biologically and toxicologically relevant to the stated purpose?
    a. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:
    b. comprehend the objective;
    c. conduct the assay;
    d. observe and measure prescribed endpoints;
    e. compile and prepare data for statistical analyses; and
    f. report the results?

3. What additional advice, if any, can be given regarding the protocol

4. Have the strengths and/or limitations of the assay been adequately addressed?

5. Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

9. Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.

Appendix B

INTEGRATED SUMMARY REPORT

Integrated Summary Report for Validation of an Estrogen Receptor Binding Assay using Rat Uterine Cytosol as Source of Receptor as a Potential Screen in the Endocrine Disruptor Screening Program Tier 1 Battery (PDF) (94 pp, 553K)

Appendix C

SUPPORTING MATERIALS

Appendix 1. Estrogen Receptor Binding (RUC) Assay Protocol
Protocol for the Estrogen Receptor Binding Assay using rat uterine cytosol (PDF) (94 pp, 444K)

Appendix 2. Background Review Document (excerpts)
Executive summary, conclusions, and recommendations from Background Review Document: Current status of test methods for detecting endocrine disruptors – in vitro estrogen receptor binding assays. ICCVAM 2002. (PDF) (65 pp, 32K)

Appendix 3. Standardized Approach for Testing Low Solubility Chemicals
Development of a standardized approach for evaluating environmental chemicals with low solubility in the estrogen receptor (ER) binding assay (PDF) (7 pp, 230K)

Appendix 4. Statistical Methods Report
Report on statistical methods for evaluating variability and setting up performance criteria for receptor binding assays (PDF) (27 pp, 802K)

Appendix 5. Second Interlaboratory Validation Study: Laboratory Summary Report
Summary report covering all three laboratories in the second interlaboratory validation study of the ER-RUC assay (PDF) (130 pp, 4.27M)

Appendix 6. Second Interlaboratory Validation Study: Lab X Data and Report
Lab X's data and report from the second interlaboratory validation study of the ER-RUC assay (PDF) (184 pp, 4.34M)

Appendix 7. Second Interlaboratory Validation Study: Lab Y Data and Report
Lab Y's data and report from the second interlaboratory validation study of the ER-RUC assay (PDF) (279 pp, 3.07M)

Appendix 8. Second Interlaboratory Validation Study: Lab Z Data and Report
Lab Z's data and report from the second interlaboratory validation study of the ER-RUC assay (PDF) (664 pp, 151M)

Appendix 9. Second Interlaboratory Validation Study: Uniform Statistical Analysis
Methods and results of the uniform statistical analysis of the second interlaboratory validation study of the ER-RUC assay (PDF) (138 pp, 1.02M)

Appendix 10. Second Interlaboratory Validation Study: Lab X Data (Re-normalized, Outliers Excluded)
Lab X data, re-normalized and fit with outliers excluded (PDF) (54 pp, 269K)

Appendix 11. Second Interlaboratory Validation Study: Lab Y Data (Re-normalized, Outliers Excluded)
Lab Y data, re-normalized and fit with outliers excluded (PDF) (149 pp, 865K)

Appendix 12. Second Interlaboratory Validation Study: Lab Z Data (Re-normalized, Outliers Excluded)
Lab Z data, re-normalized and fit with outliers excluded (PDF) (93 pp, 660K)

Appendix 13. Second Interlaboratory Validation Study: Superimposed Graphs of Acceptable Runs

Superimposed graphs of acceptable runs for reference standard (estradiol), weak positive (norethynodrel), and test chemicals, by laboratory (PDF) (29 pp, 1.15M)